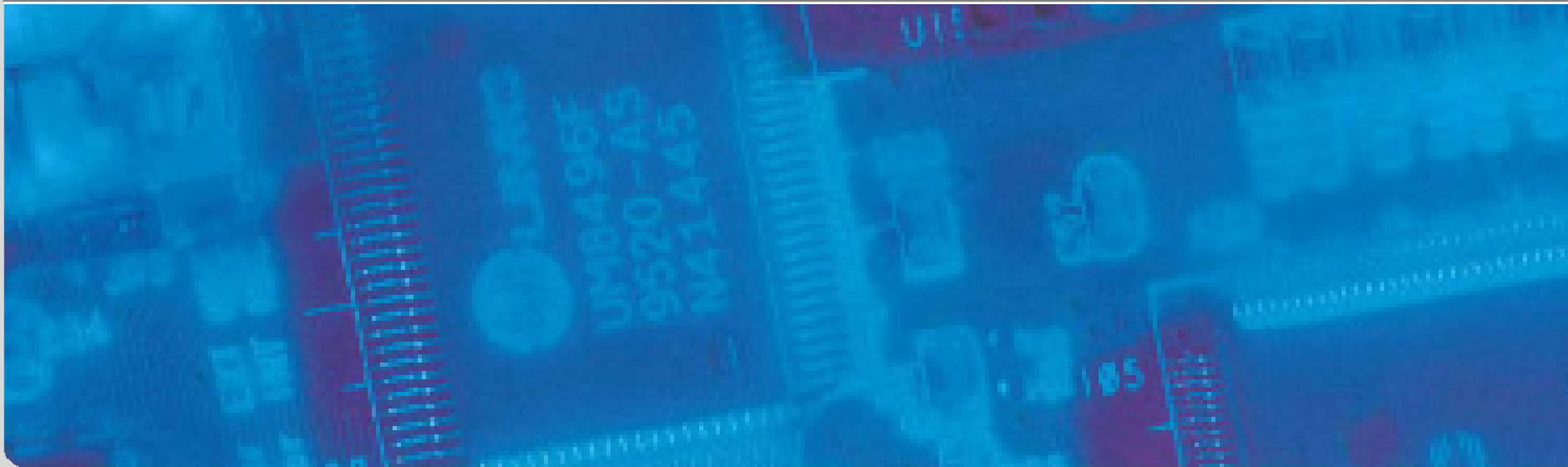


Reliability of On-Chip Systems - A Thermal Perspective -

by J. Henkel

CES – Chair for Embedded Systems



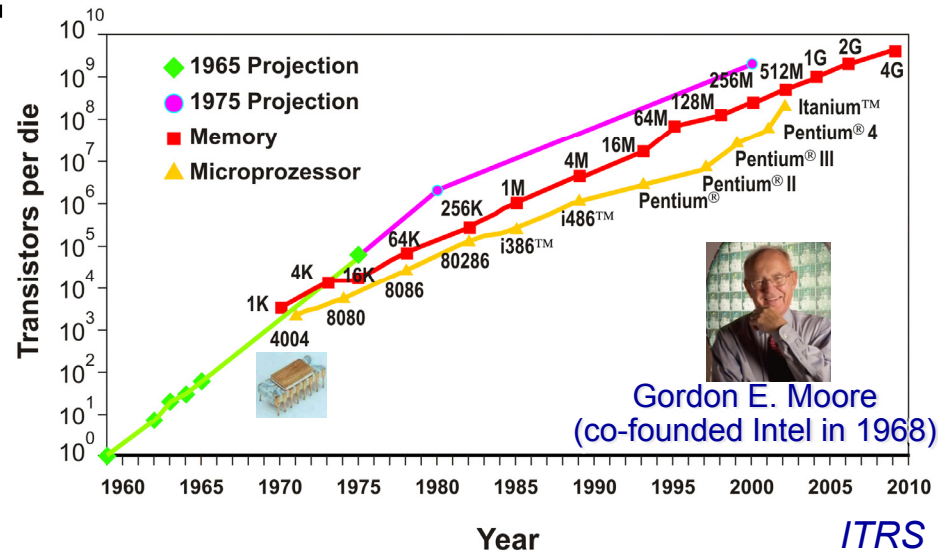
Outline

- Dependability Problems
- Dependability and Thermal Issues
- Counter Measures
- Thermal Management

Outline

- Dependability Problems
- Dependability and Thermal Issues
- Counter Measures
- Thermal Management

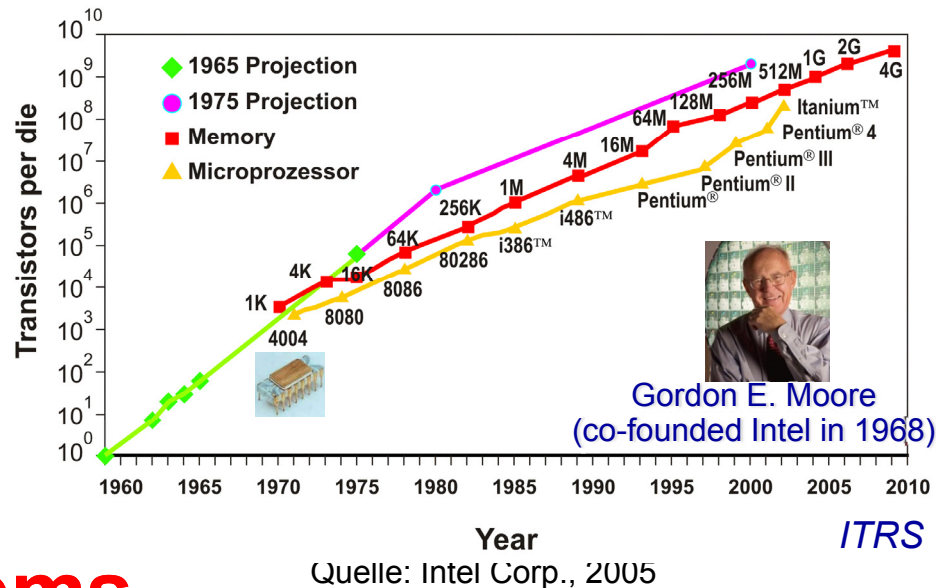
In the Past ...



■ ... Moore's Law provided a win-win situation:

- Smaller feature size
- Higher integration density, more functionality
- Lower power consumption
- Higher speed (performance)
- Less cost (per-transistor costs)
- ...

In the Future ...



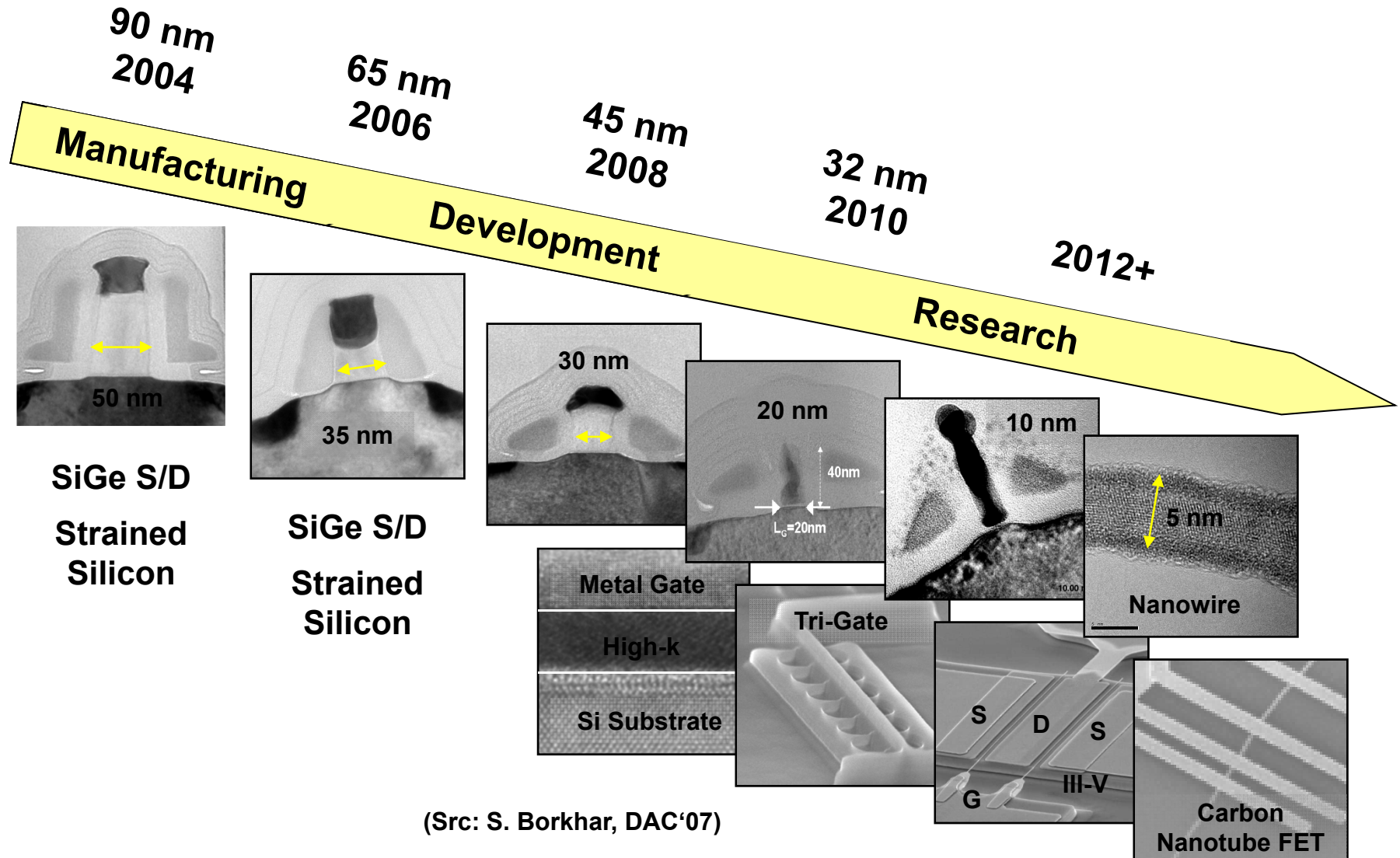
□ ... Problems

- Complexity: In 2017 100 Billion Transistors on chip
- Productivity gap
- Thermal problems
- Increasing relevance of aging effects
- Manufacturing defects, process variation
- Stochastic effects since physical limits are reached
- Decreasing yield



Reliability

Technology Nodes



Variabilities

■ Variability of transistor structures

- Channel Length
- Isolators thickness (gate oxid) gate <-> transistor channel
- Randomized Dopant Fluctuations (RDF) -> Threshold voltage

=> Decreasing mobility

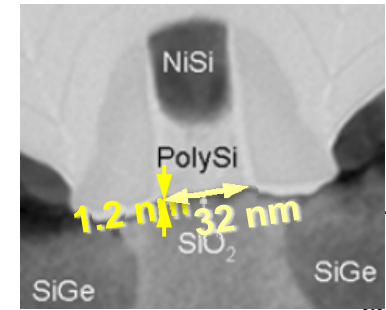
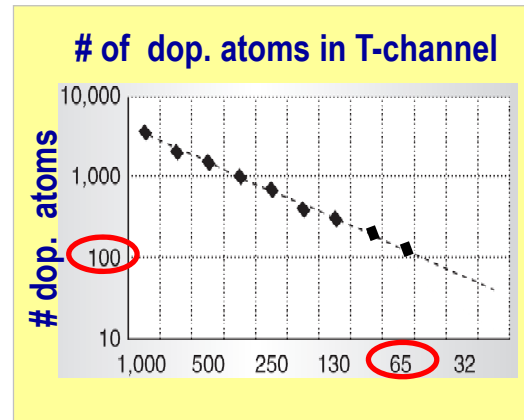
=> Increasing leakage

■ Counter Measures

- Strained Silicon Engineering
 - Strain channel to increase mobility
- „High-K“ materials for gate isolation (e.g. Hafnium)
 - May increase aging

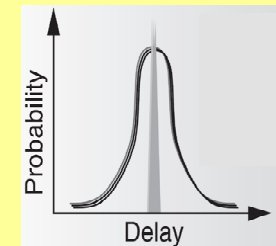


...

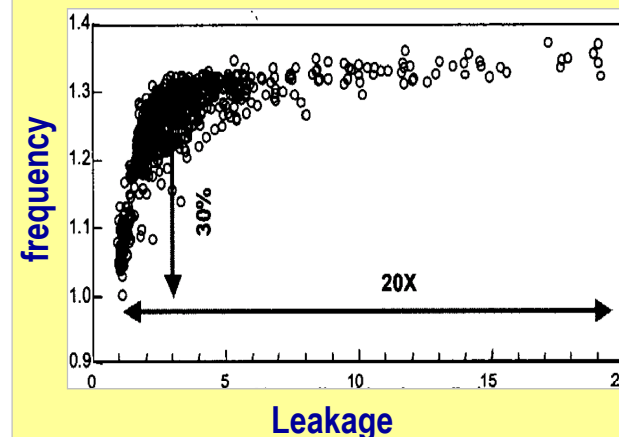


Quelle: Intel, 65 nm

latency of an Inverter



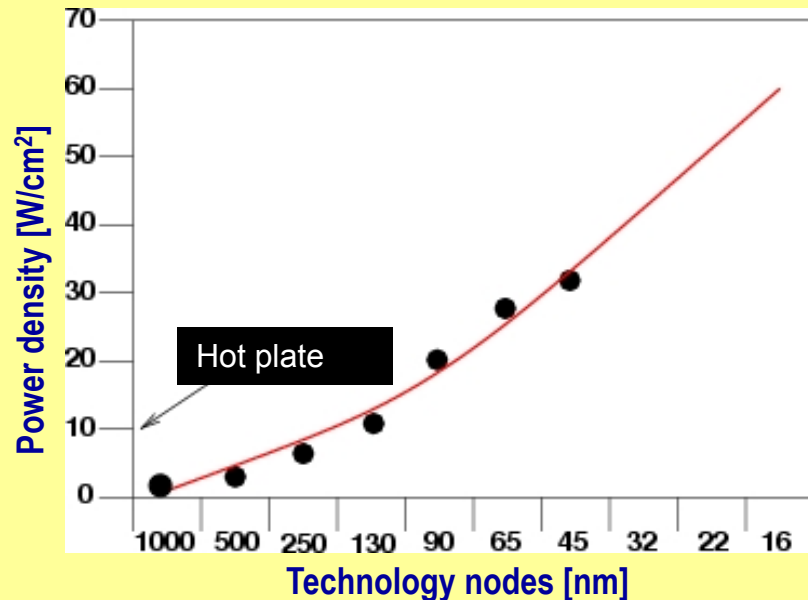
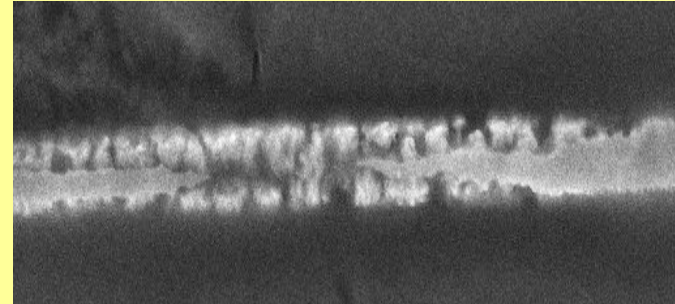
Variations for processors an single die



Aging Effects

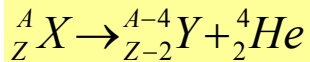
- Elektromigration (EM)
- Stress Migration
- Time-dependent Dielectric Breakdown
- Dependent upon operating temperatur
- Through technology scaling
 - Increasing frequency
 - Increasing power dissipation per area, volume

Wire affected by electro migration

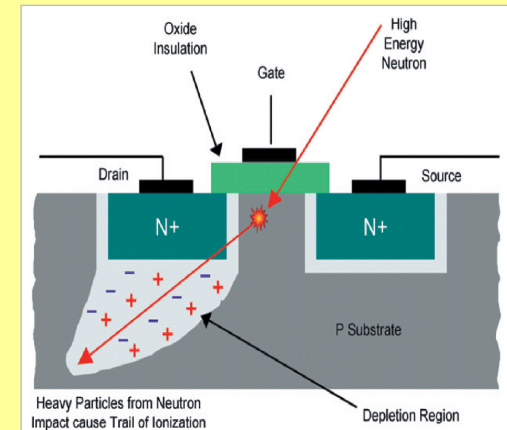


Increasing Susceptibility to Soft Errors

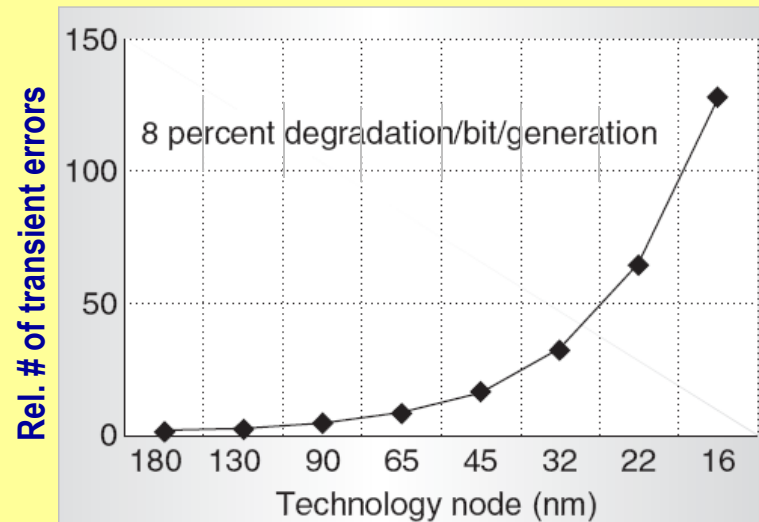
- Ionizing rays may change charge concentration
 - (like He^{2+})
=> may lead to bit flips
- α -rays
 - Radioactive decomposition of non-pure chip material
- Cosmic rays (particularly neutrons)
- accelerated through technology advancements
 - Low voltage and capacitances
 - Representation of bits through smaller and smaller charges



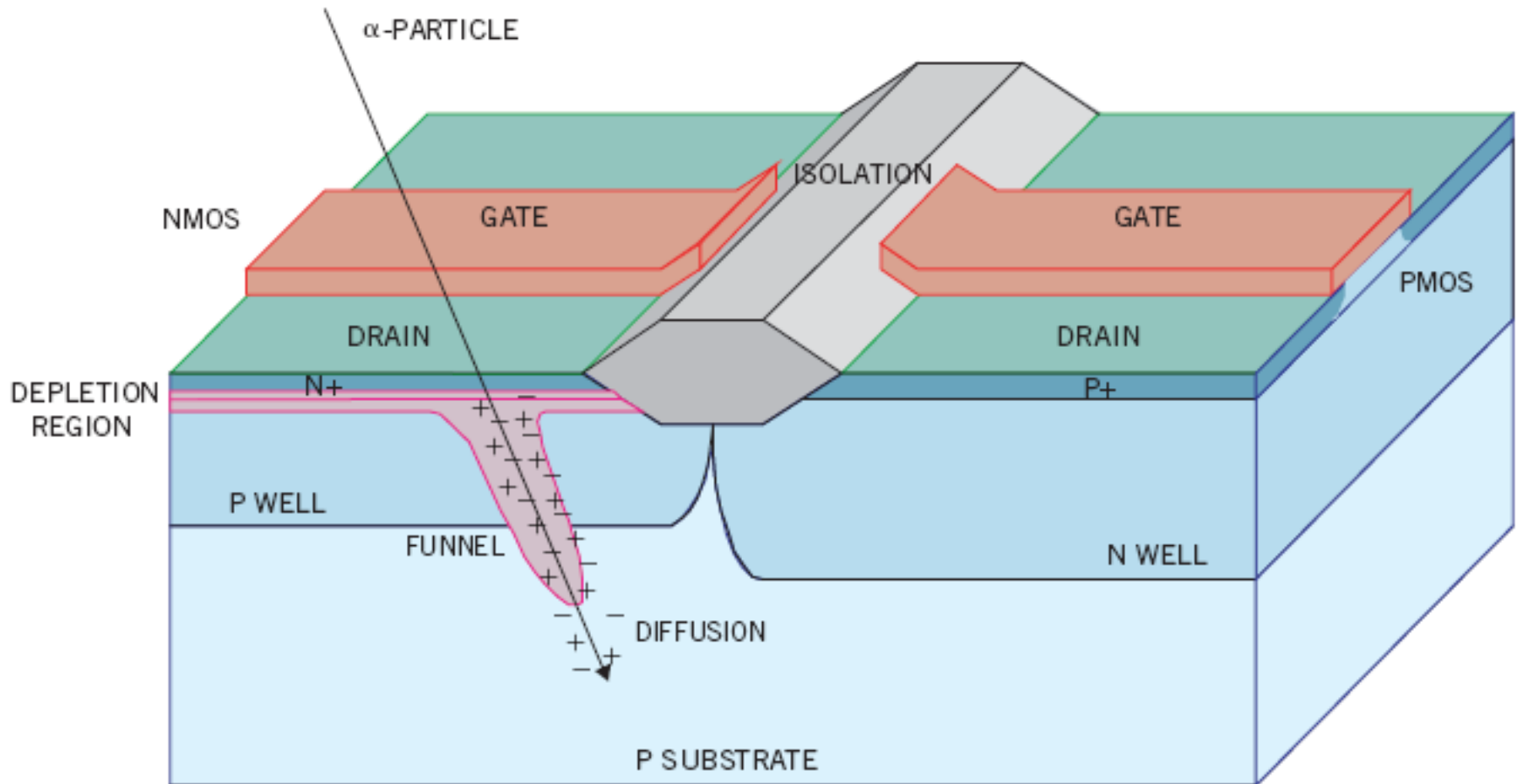
Transient errors through neutrons



Transient errors



Particle strikes: causing soft errors



(Src: R. Mastipuram: Cypress Semiconductor @ EDN, Design Feature'04
Soft errors' impact on system reliability)

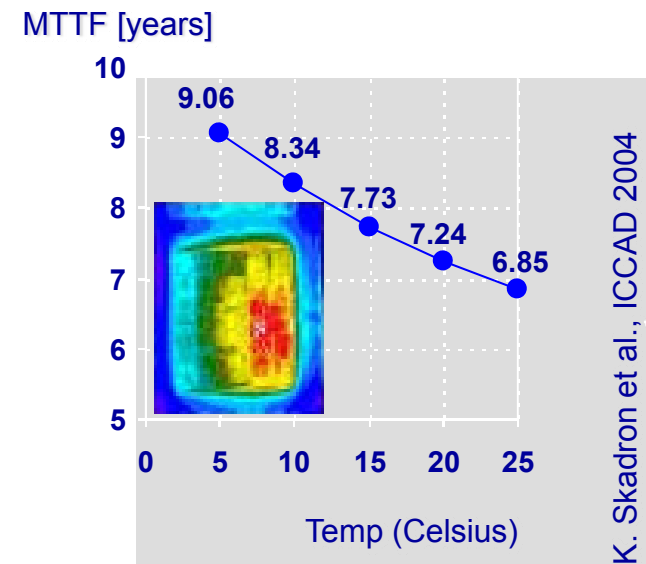
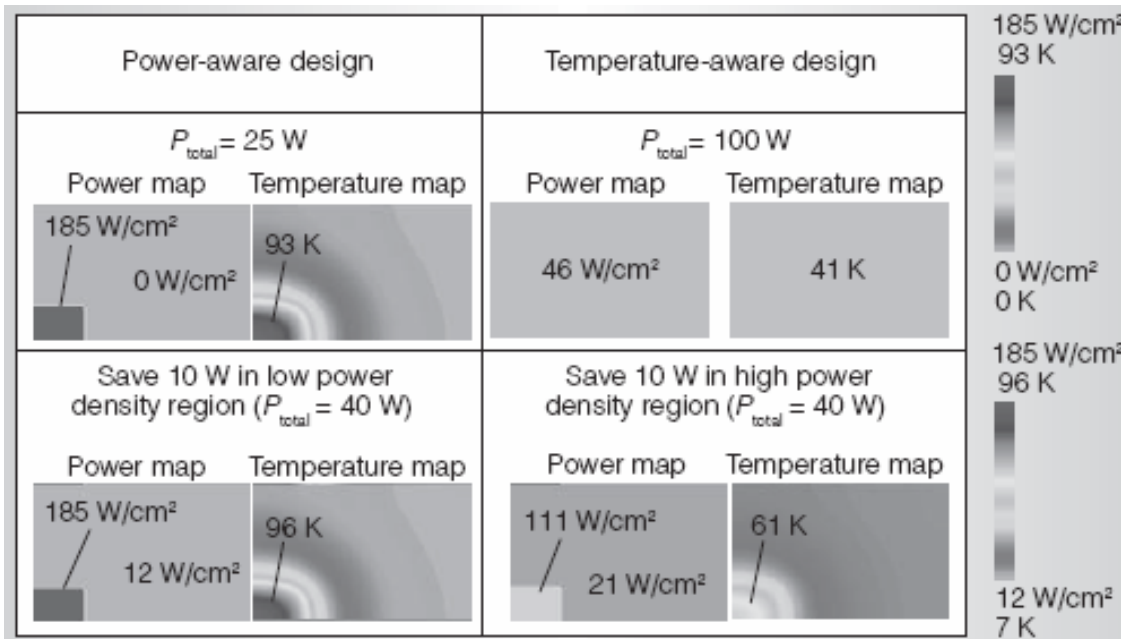
Outline

- Dependability Problems
- Dependability and Thermal Issues
- Counter Measures
- Thermal Management

Heat Remains a Problem

“Circuit heat generation is the main limiting factor for scaling of device speed and switch circuit density”

*By Jeff Welser, Director SRC Nanoelectronics Research Initiative, IBM,
Opening Keynote Address ICCAD 2007*



(Src: K. Skadron: Low-Power Design and Temperature Management;
IEEE Micro, Vol. 27, No. 6, 2007)

From Power to Temperature

- Heat is thermal energy [Joule]
- Heat transfer Q [Joule/s]
- Heat flux is heat transfer rate through given surface area
- Thermal capacity C :

$$C = \frac{\Delta Q}{\Delta T}$$

- Temperature T reflects the amount of heat energy given a certain material

Specific heat capacity c

Material	$c=C/\text{dm}^3$
Silica	1.55
Cu	3.45
H ₂ O	4.17
Air	0.0012

From Power to Temperature (cont'd)

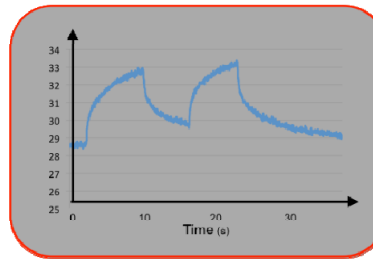
- Basic temperature equation:

$$C \frac{dT}{dt} = -Q + P \qquad T(t_1) = T_0 + \frac{1}{C} \int_{t_0}^{t_1} -Q(t) + P(t) dt$$

where Q is the heat dissipation rate.

$$T(t) = T_0 - (T_{SS} - T_0) e^{-\frac{t}{h}}$$

Heating



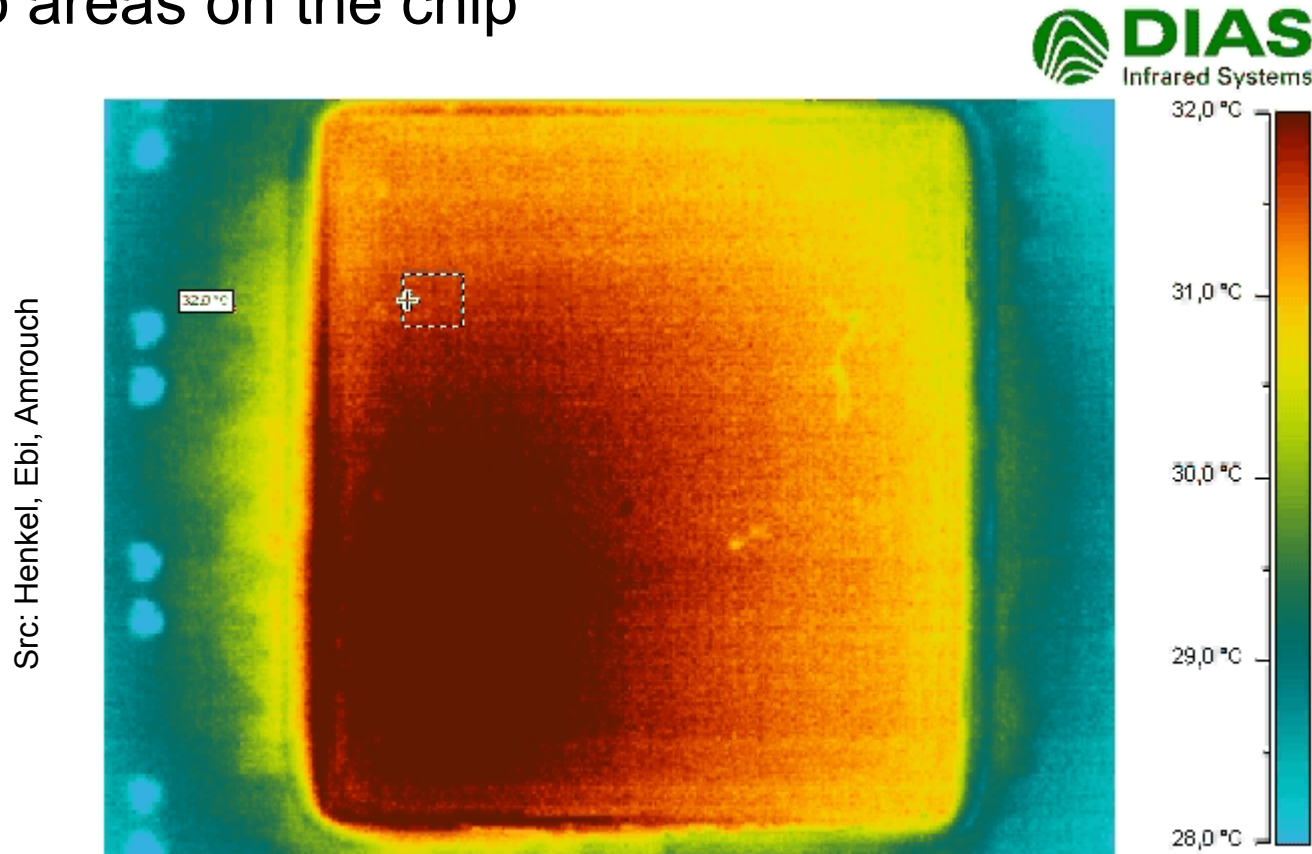
$$T(t) = T_0 + (T_0 - T_A) e^{-\frac{t}{c}}$$

Cooling

- T_{SS} is the steady state temperature the system will asymptotically reach with current power configuration
- Ambient temperature T_A is minimum reachable temperature

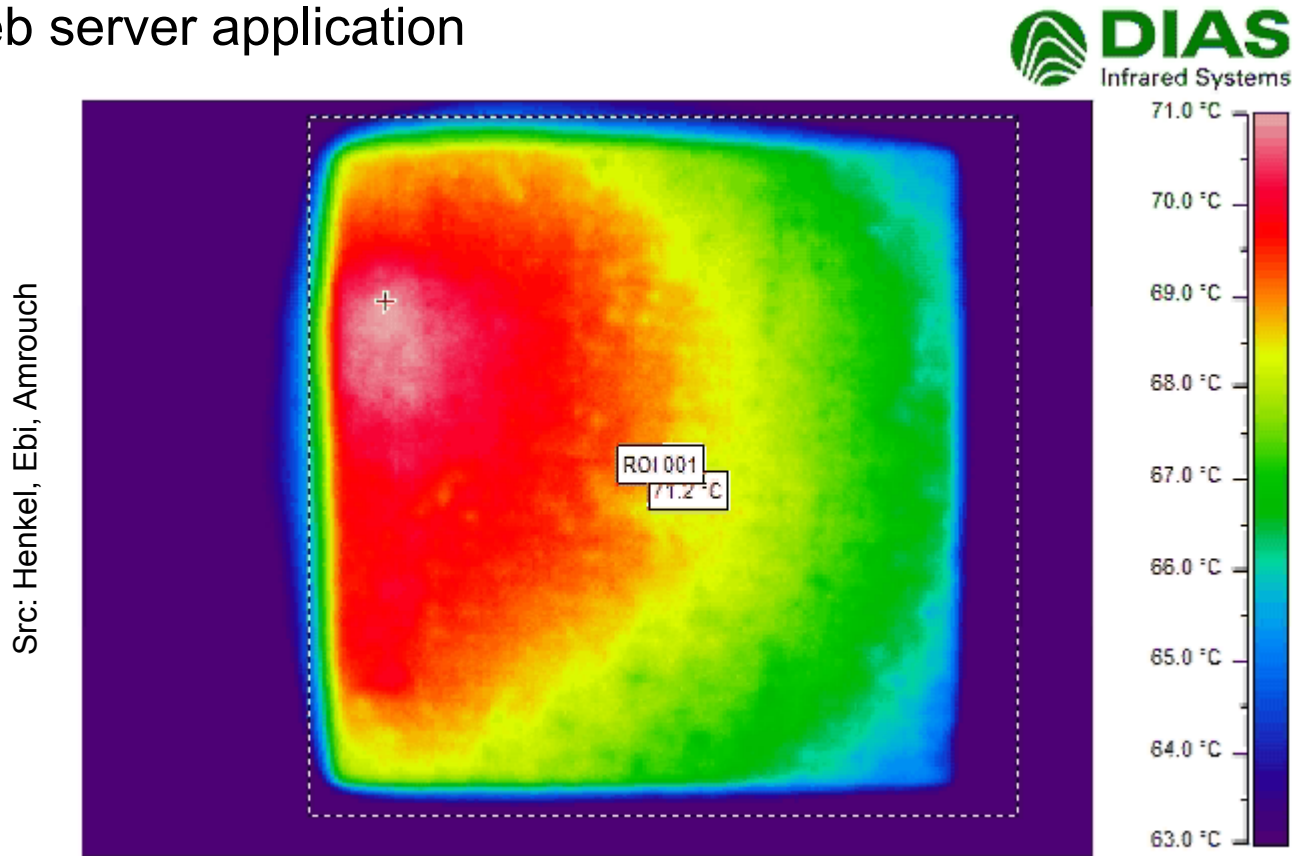
Thermal Distribution and Dynamics

- Example showing localized computation switching between two areas on the chip



Thermal Distribution and Dynamics (cont'd)

- Example: Xilinx Virtex 5 running a web server application



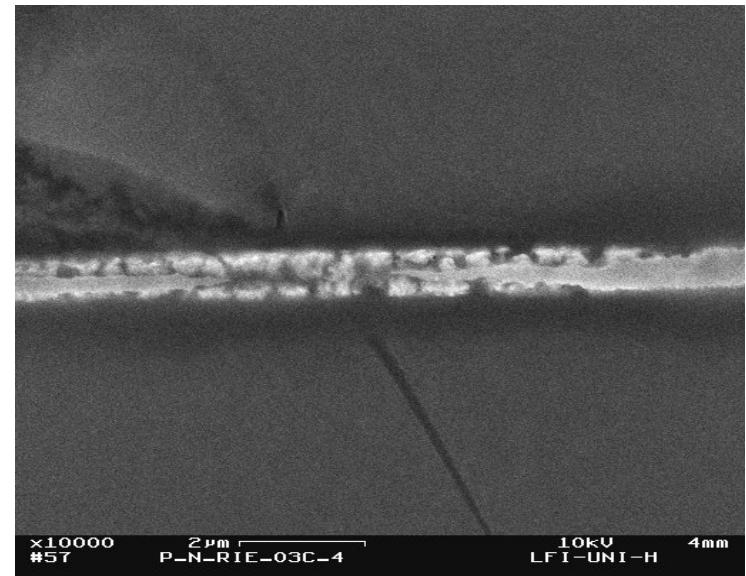
From Temperature to Reliability

- For instance: Electromigration:
- directly linked to temperature
 - Basic *Mean time to failure* modeled by Black's Equation:

$$MTTF = Aj^{-n} e^{\left(\frac{Q}{kT}\right)}$$

- *MTTF* decreases exponentially with temperature

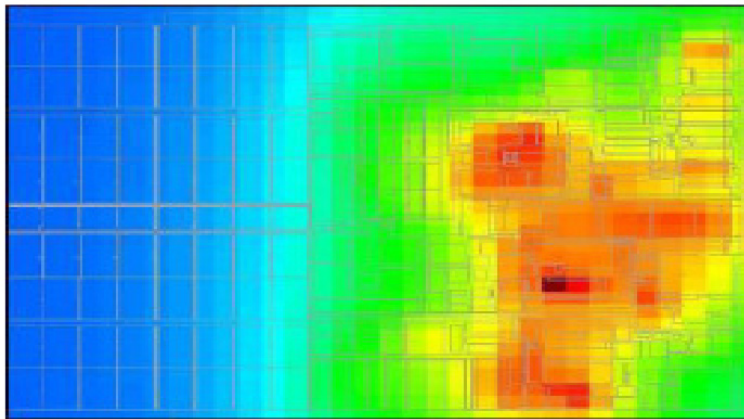
→ **Goal**: reduce peak temperatures



[wikipedia]

From Temperature to Reliability

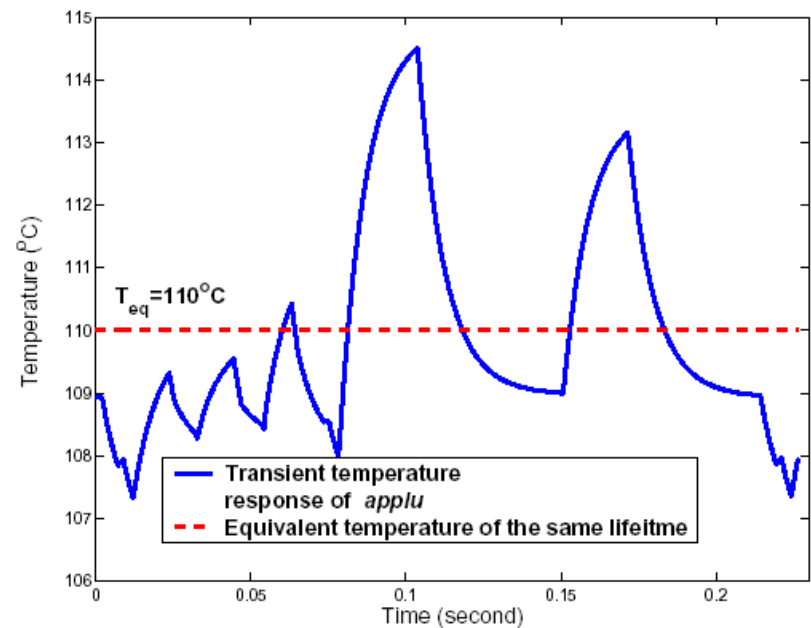
- *MTTF* also affected by thermal gradients



Spatial gradients

Simulated Thermal map Pentium M

[L.Finkelstein, Intel 2005]



Temporal gradients

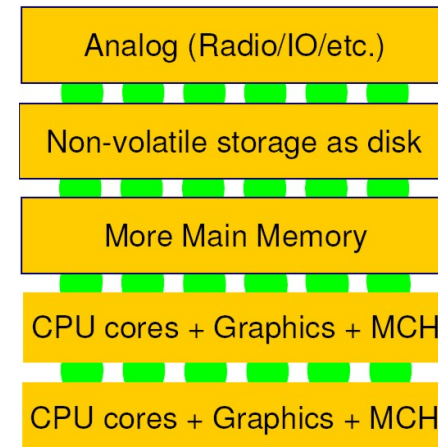
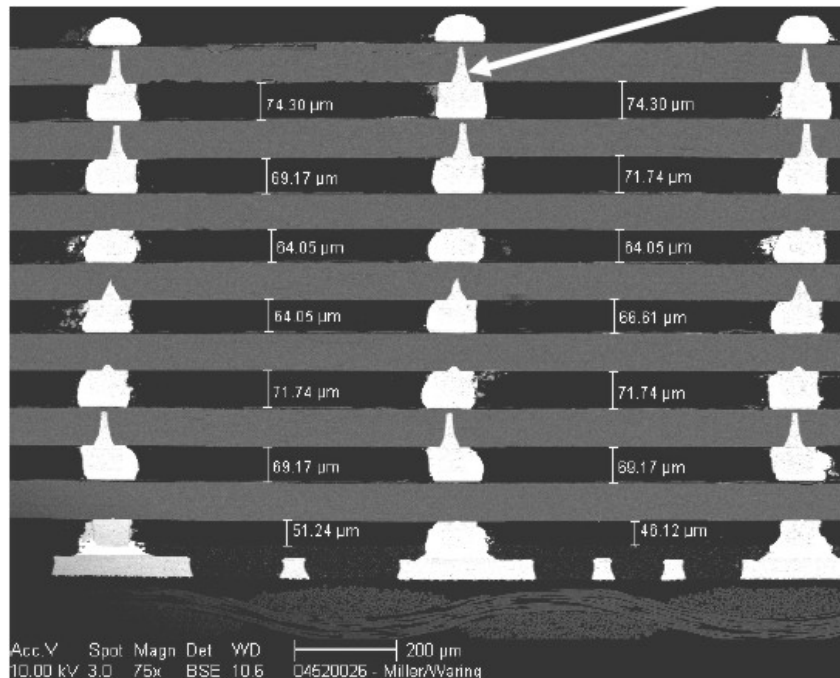
[K. Skadron, 2005]

- → **Goal**: balance temperatures

Thermal/Heat Problems in 3D

- 3-D chips especially problematic

3-D structures

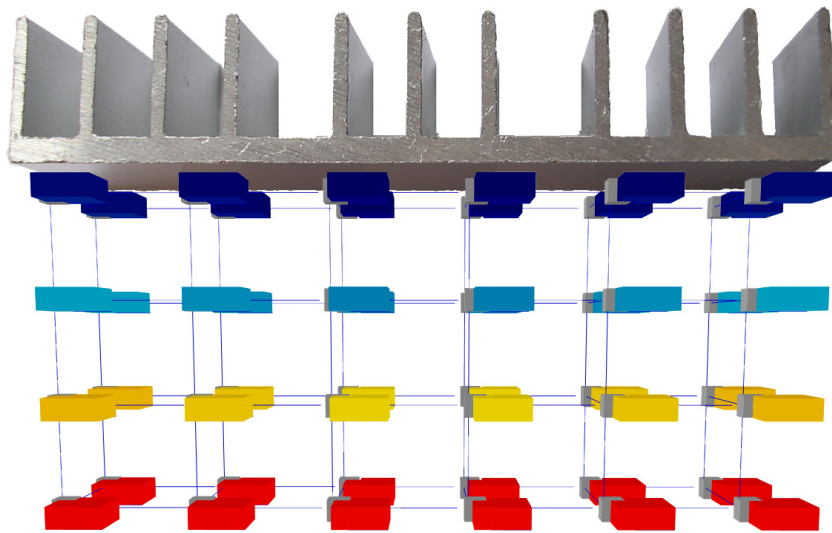


(Src: Y. Xie, PennState)

Thermal/Heat Problems in 3D Architectures

■ Problem: vertical heat flow

- Only one layer directly interfaces with the heat sink
- Heat needs to dissipate through multiple layers



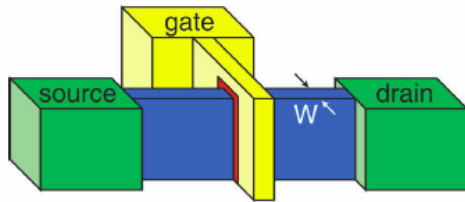
- The heat sink is located on top of the chip
- Hot cores distant to the heat sink dissipate their heat through other layers
- Silicon has a *low thermal conductivity!*
 - $150 \text{ W/(m}^{\circ}\text{K)}$ (Silicon)
 - $401 \text{ W/(m}^{\circ}\text{K)}$ (Copper)

Outline

- Dependability Problems
- Dependability and Thermal Issues
- **Counter Measures**
- Thermal Management

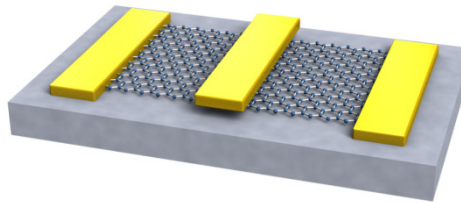
Counter Measures at Device Level

FinFET-Transistor

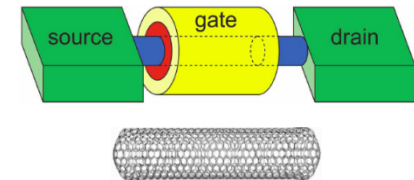


Idea: reduce channel thickness
But: reduced mobility

Graphene-Transistor

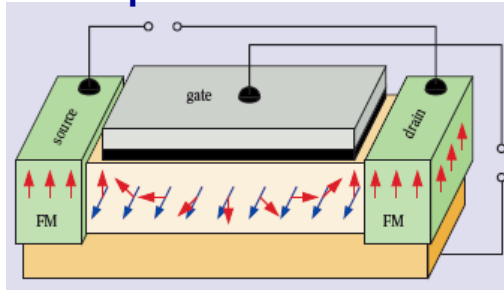


CNFET-Transistor



Idea: combine high mobility and thin channel width
But: problems in placement and structural growth

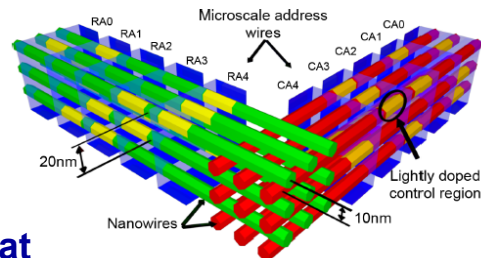
Spin-Transistor



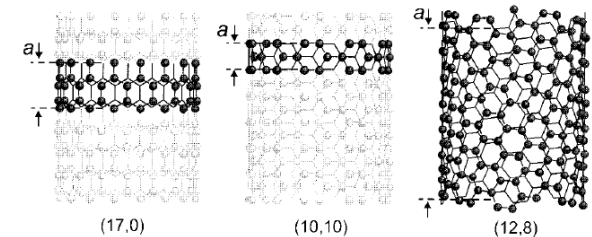
Injection of spin-polarized electrons at source V-gate affects spin trace electron current only when electron spin parallel to drain-spin

Idea: low power dissipation
But: hard to control => high error rates

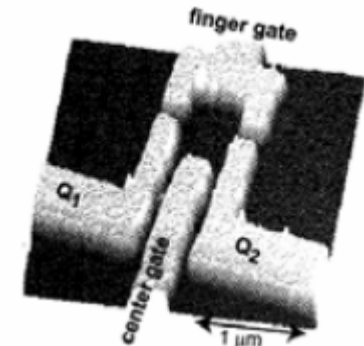
NanoPLA block and 3D Interconnect



Source: DeHon



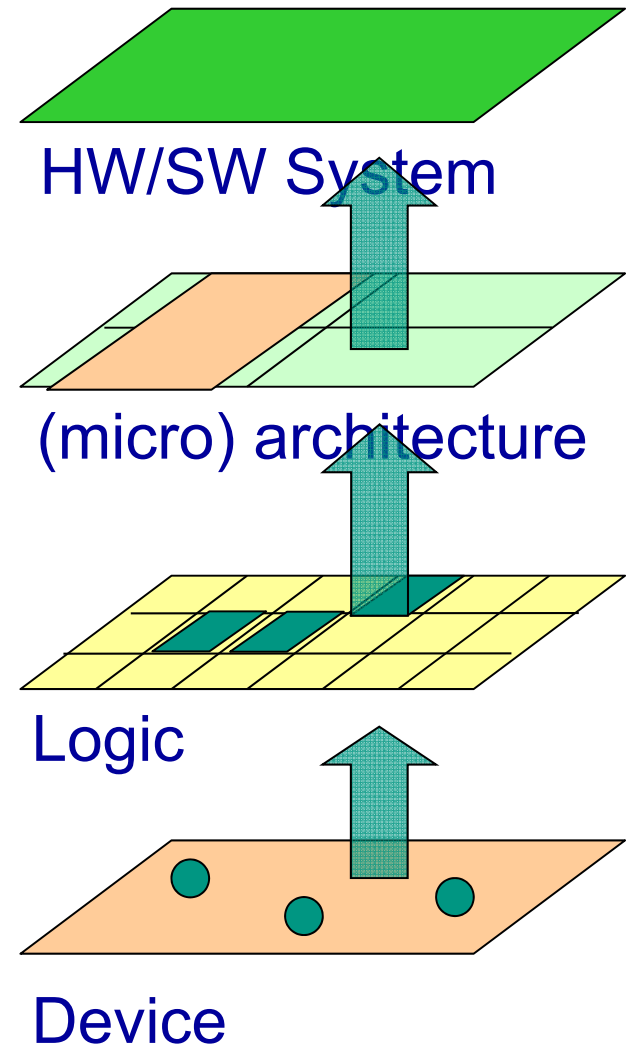
Single-Electron Transistor



A Spectrum of Solutions

- ❑ Near and medium term solutions:
 - ❑ Massively parallel, modularity (cells, blocks)
 - ❑ Regularity (grid processing, cellular arrays)
 - ❑ Locally connected (near-neighbor connections, crossbar)
 - ❑ Higher functionality (multiple valued logic, threshold logic)
 - ❑ Adaptivity through Reconfigurability
 - ❑ Asynchronous (including GALS)
 - ❑ Fault-tolerance (noise immune,, redundant, self-testing, self-correcting)
 - ❑ Defect-tolerant (reconfigurable)
 - ❑ Redundant, adaptive (self-adaptive, self-organizing, evolvable)
 - ❑ **Bio-inspired, autonomous computing etc.**
 - ❑ Nanophotonic (optical communication, GOLE)
 - ❑ Nanofluidic
 - ❑ 3D interconnects
 - ❑ Probabilistic (algorithms, encoding, communication)
- ❑ Long term solutions:
 - ❑ molecular, quantum
 - ❑ quantum-dot cellular automata
 - ❑ Adiabatic / reversible
 - ❑ ...

(Src: M. Huebner, KIT)



Idea: use Principles of Bio-inspired/Autonomous Computing

■ Organic Computer Systems

- will possess lifelike properties.
- will consist of autonomic and cooperating sub systems and will work, as much as possible, in a self-organized way.
- will adapt to user needs,
- will be controlled by objectives (“goal-driven”),

■ Self-organization allows for adaptive and context aware behavior:

■ Self-X

(Src: H. Schmeck, KIT)

- self-configuring
- self-optimizing, self-adapting
- self-healing
- self-protecting
- ...



=> Beneficial for reliability

Self-Organization

- Under appropriate conditions the **collaboration of simple agents may** produce highly complex, adaptive systems
- **No** necessity for **central control**
- Examples: (Src: H. Schmeck, Uni Karlsruhe)
 - Termite / Ant colonies
 - Swarms of bees
 - Economy
 - Traffic
 - Internet
- **Idea:** Complexity management by self-organization
- But: Who is managing/controlling self-organisation?

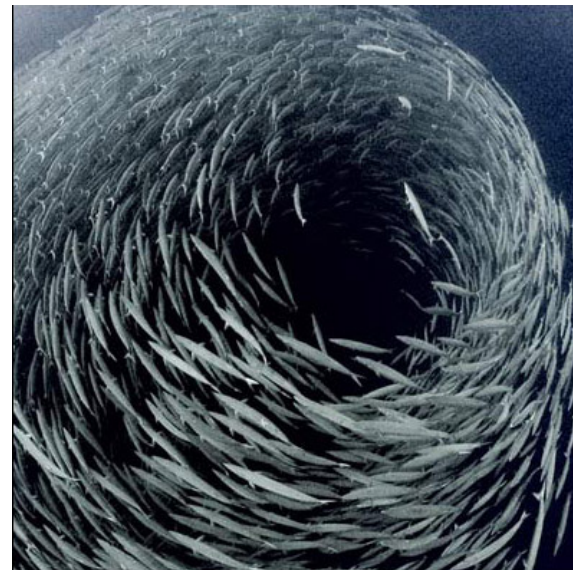


Emergent Phenomena

- Local interaction may lead to entirely new global properties.

„The whole is more than the sum of its parts!“

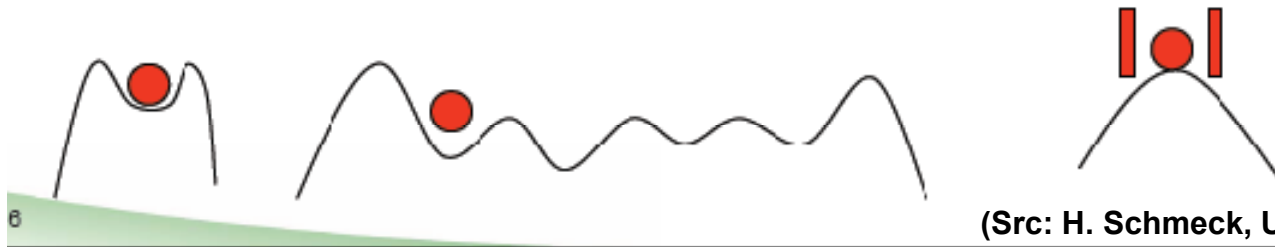
- Emergent effects may be desired or undesired
 - How can we generate positive emergence?
 - How can we prevent negative emergence?
- Examples:
 - „green wave“ at traffic lights
 - deadlock / lifelock
 - ant roads
- Can we use emergent Phenomena for computing systems?



(Src: H. Schmeck, Uni Karlsruhe)

Challenges in Bio-inspired/Autonomous Computing

- Provide systems with sufficiently large degrees of freedom for adapting to different requirements.
- Systems have to be aware of
 - what type of service they can provide,
 - what type of service they need from others,
 - what the current environment wants to get done.
- Systems should have a “desire” to be active (□□incentives?).
- Systems should be robust with respect to external changes
- Systems should react flexibly to changing external constraints
- There will be a need for “controlled self-organization”.



(Src: H. Schmeck, Uni Karlsruhe)

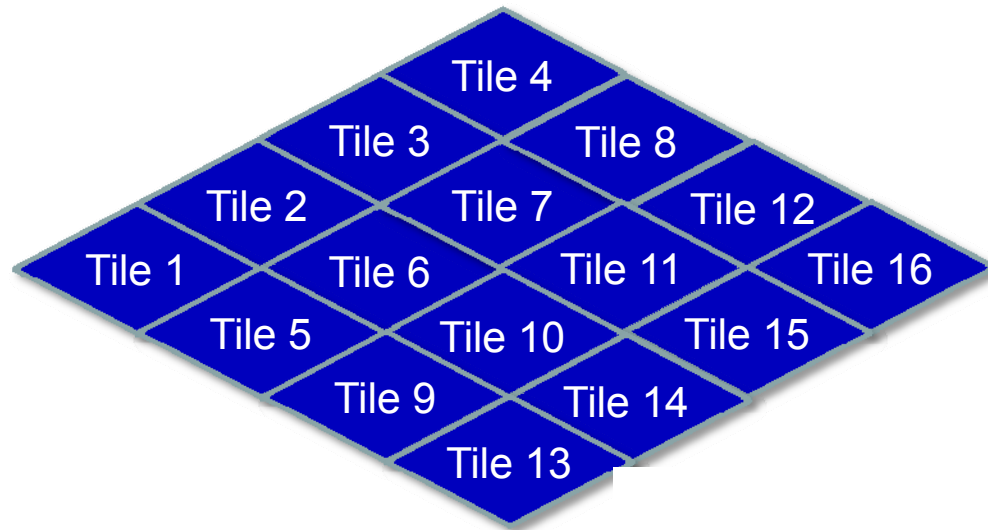
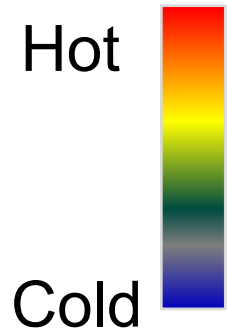
Bio-inspired/Autonomous Computing

- Self-organizing computing systems are becoming a key topic for academic and industrial research.
- So, what do we need?
 - Nature inspired methods, Artificial Life:
 - Evolutionary Algorithms, Ant Colony Optimization, Swarm intelligence
 - **Multi-Agent systems**
 - Cognitive systems
 - Learning
 - Observer/Controller-Architectures
 - Results from control theory (model predictive control??)
 - Reconfigurable computing systems (Src: H. Schmeck, Uni Karlsruhe)
 - ...
- **Organic/Autonomous Computing may help to build more reliable systems!**

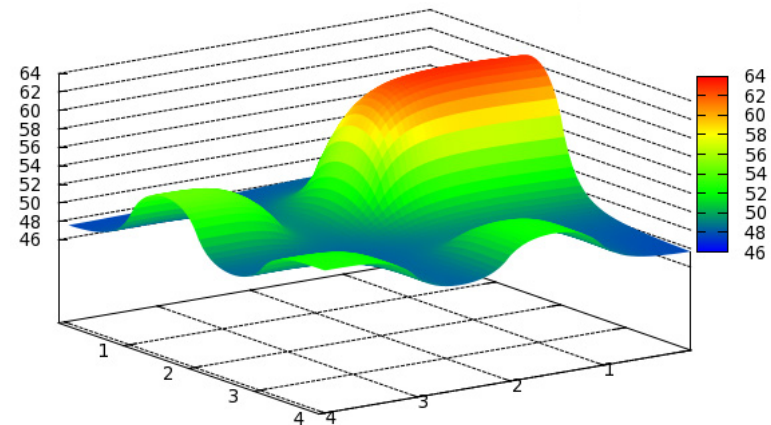
Outline

- Dependability Problems
- Dependability and Thermal Issues
- Counter Measures
- Thermal Management
 - Using principles of self-organization:
 - Scalability
 - Proactivity
 - No single point of failure

Motivation



Spreading tasks
throughout the chip
reduces thermal hotspots



→ Thermal hotspots!

Motivation

Open Problem

Thermal hotspots in multi/many-core architectures

Possible Solution:

Dynamic Thermal Management (DTM)

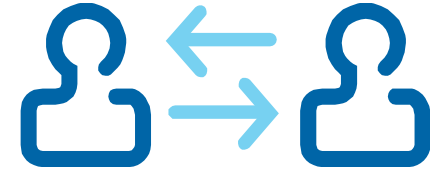
- Design-time techniques may not predict the behavior a priori
- Runtime application mapping algorithm may be used to homogeneous thermal distribution

What Properties are Required?

- Scalability
- Proactive behavior
- Light-weight in terms of hardware/software

Idea: use Agent-Based System

“An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors”



[Russell & Norvig, Artificial Intelligence: A modern Approach]

■ Desired properties of agent

- **Situated** \leftrightarrow Software/hardware entity in each tile
- **Scalable** \leftrightarrow Agents act locally
- **Proactive** \leftrightarrow Triggered before threshold is reached
- **Social** \leftrightarrow May negotiate with their neighbors
- **Reactive** \leftrightarrow React to outside stimuli (i.e. to thermal sensors)
- **Light-weight** \leftrightarrow Require small memory/computation footprint

Agent-Based System

Approach

- Economic policy to achieve proactive behavior
- Distributed approach for mapping using agents

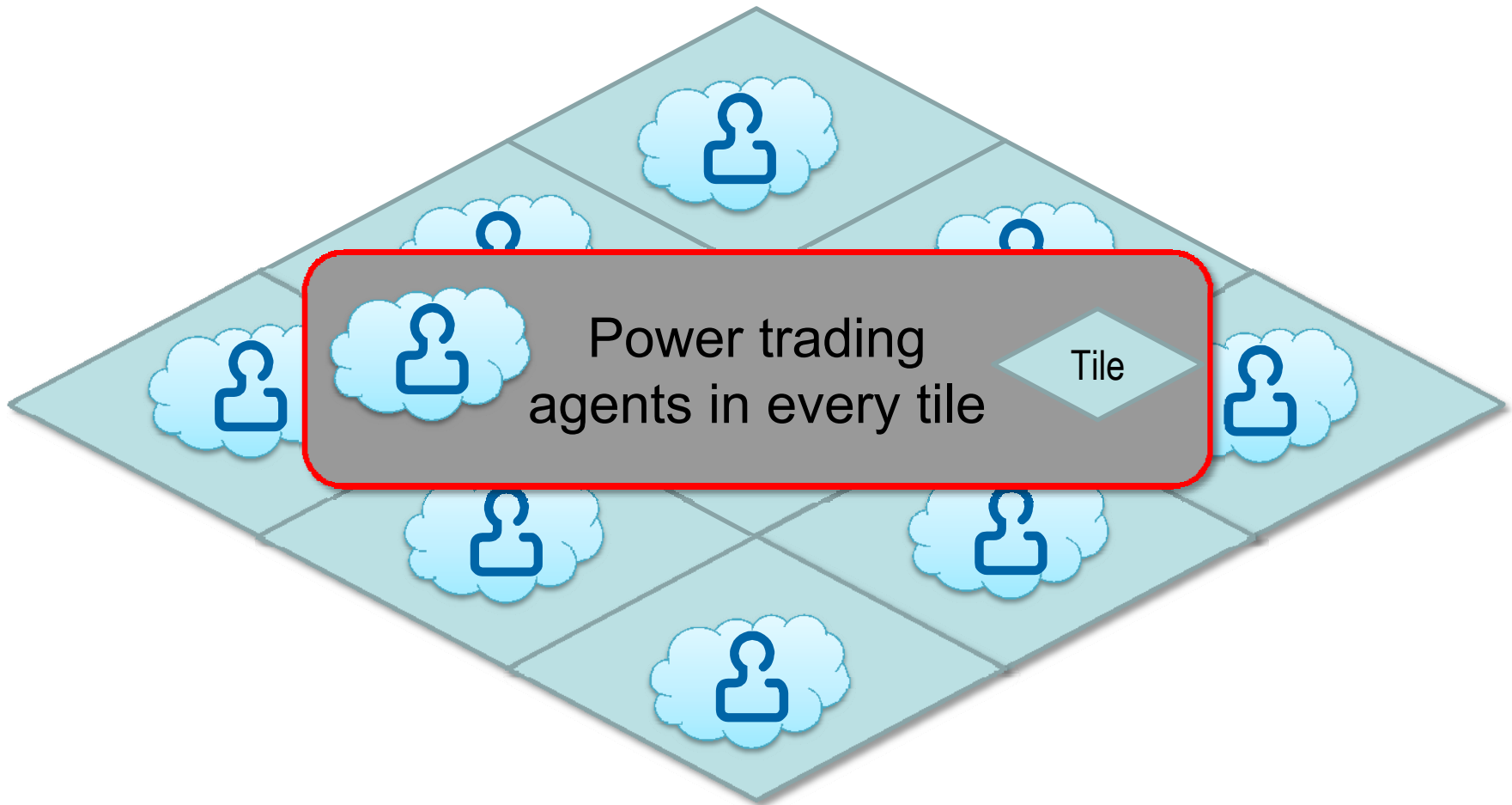
Power Trading Agents

- Economic policy (supply/demand) to achieve proactive behavior
- HW/SW implementation
- Situated in each tile

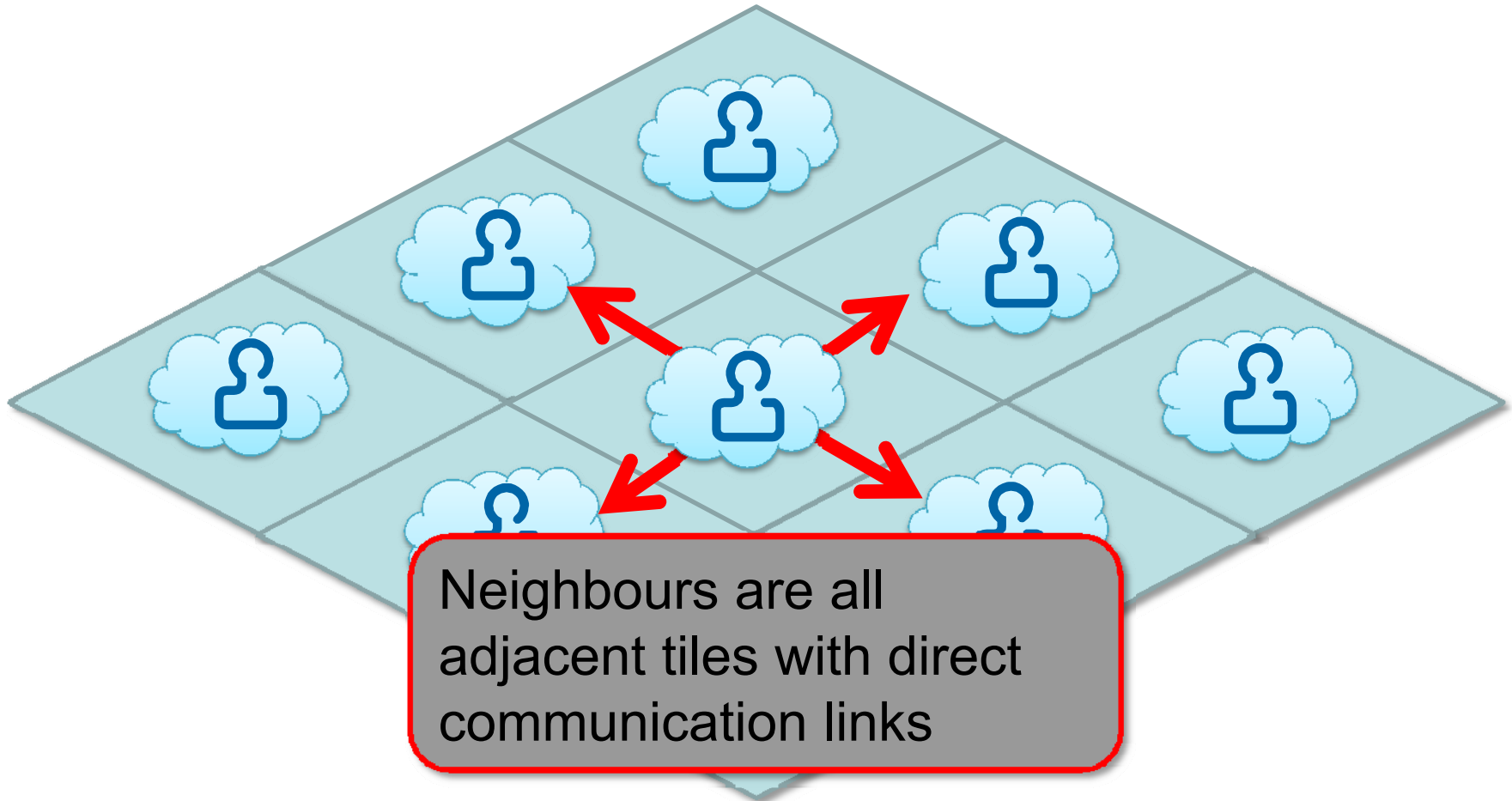
Mapping Agents

- Distributed approach for mapping using agents
- SW implementation
- Responsible for a region of neighboring tiles
- Can be migrated

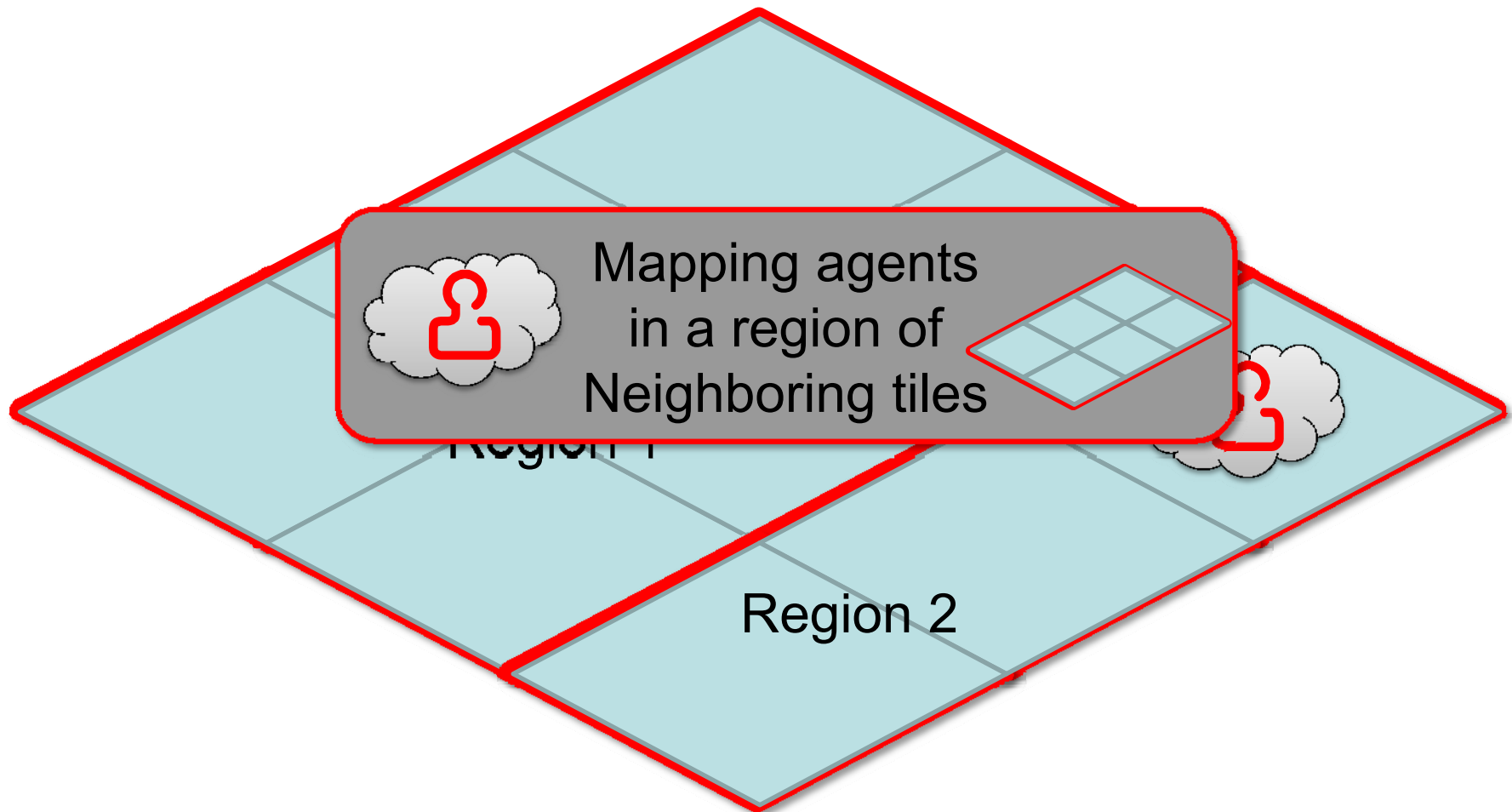
Agent-Based System



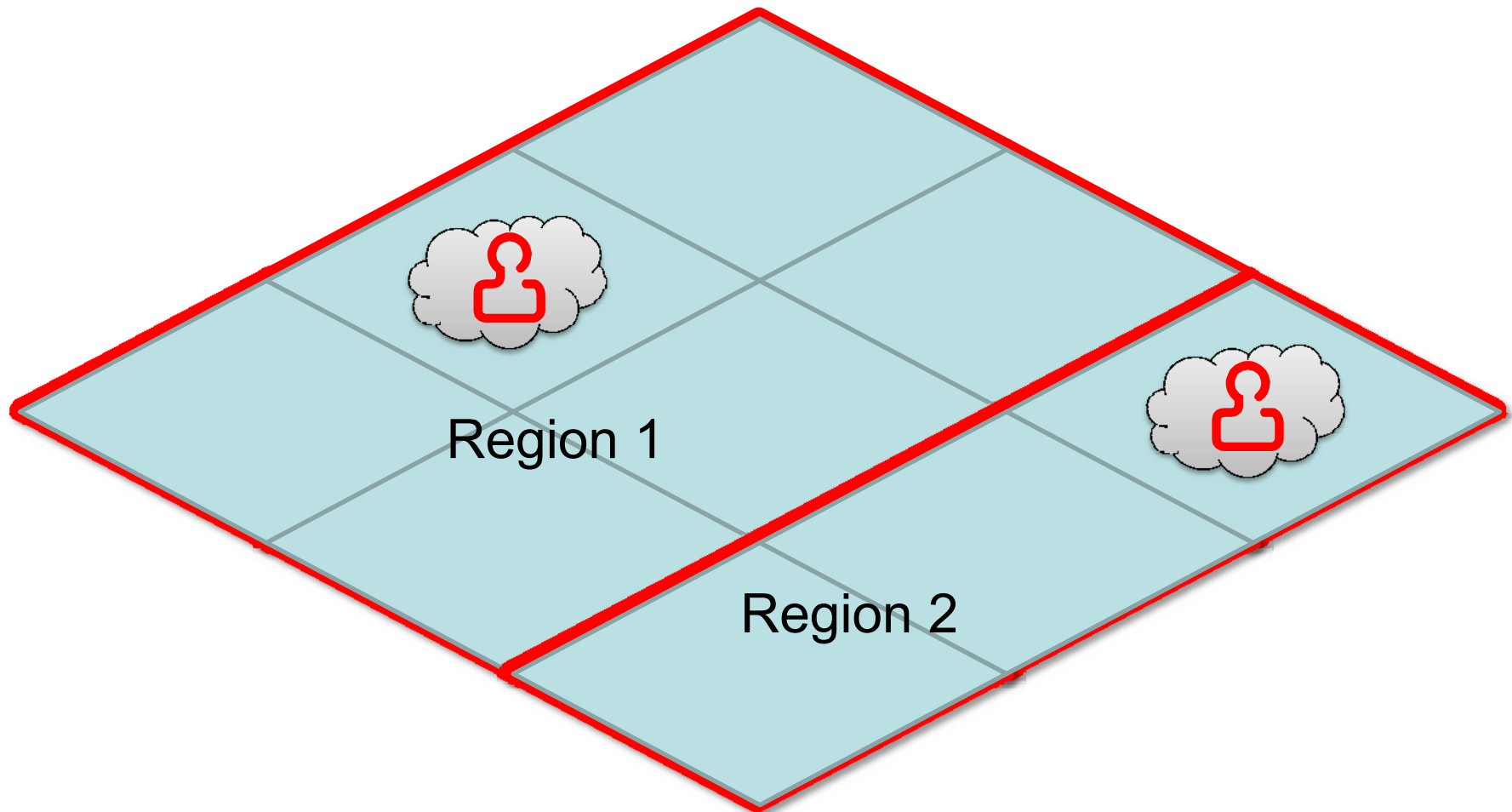
Agent-Based System



Agent-Based System



Agent-Based System



Trading Units



Number of power units
→ frequency f , voltage V
tasks can be run

Units traded between agents are **power units**

Used power units:
used to run tasks => refers to a
certain voltage/frequency setting

Free power units:
can be freely traded among agents

Assumptions, Parameters

Tasks have fixed deadline

**A worst case of the execution time (WCET)
is known**

- Minimum frequency is set for a tile**
- This results in the number of 'used' power units**

**Task migration is system dependent
(measured around 100K cycles
(saving, transferring, loading task context))**

Trading Rules

- *Buy and sell* **incentives** express an agent's "desire" to acquire/give up power units => based on "**supply/demand**" **like in an economy**

Trading Rules

- *Buy and sell incentives* express an agent's "desire" to acquire/give up power units => based on "supply/demand" like in an economy.
- Incentive to Sell: $sell = w_{u,s} \cdot used + w_{f,s} \cdot free$
- Incentive to Buy: $buy = w_{u,b} \cdot used - w_{f,b} \cdot free + \gamma$

Trading Rules

- Buy and sell **incentives** express an agent's "desire" to acquire/give up power units => based on "supply/demand" like in an economy.

- Incentive to Sell: $sell = w_{u,s} \cdot used + w_{f,s} \cdot free$

- Incentive to Buy: $buy = w_{u,b} \cdot used + w_{f,b} \cdot free + \gamma$

Weights and γ are dependant on processor type, total amount of power units, and number of tiles

Logically: used units => **demand**
free units => **supply**

Trading Rules

- *Buy* and *sell* **incentives** express an agent's "desire" to acquire/give up power units => based on **"supply/demand" like in an economy**
- Incentive to *Sell*: $sell = w_{u,s} \cdot used + w_{f,s} \cdot free$
- Incentive to *Buy*: $buy = w_{u,b} \cdot used - w_{f,b} \cdot free + \gamma$
 $buy = buy - a_b \cdot temp$
- Thermal penalties:
 $sell = sell - a_s \cdot temp$

temp is temperature
above threshold T_0

Trading Rules

- Buy and sell **incentives** express an agent's "desire" to acquire/give up power units => based on "supply/demand" like in an economy

- Incentive to Sell : $sell = w_{u,s} \cdot used + w_{f,s} \cdot free$

- Inc Agent of tile n sells to neighbor i if: $sell_n + \gamma$

$$(sell_n - buy_n) - (sell_i - buy_i) > \varepsilon$$

- The

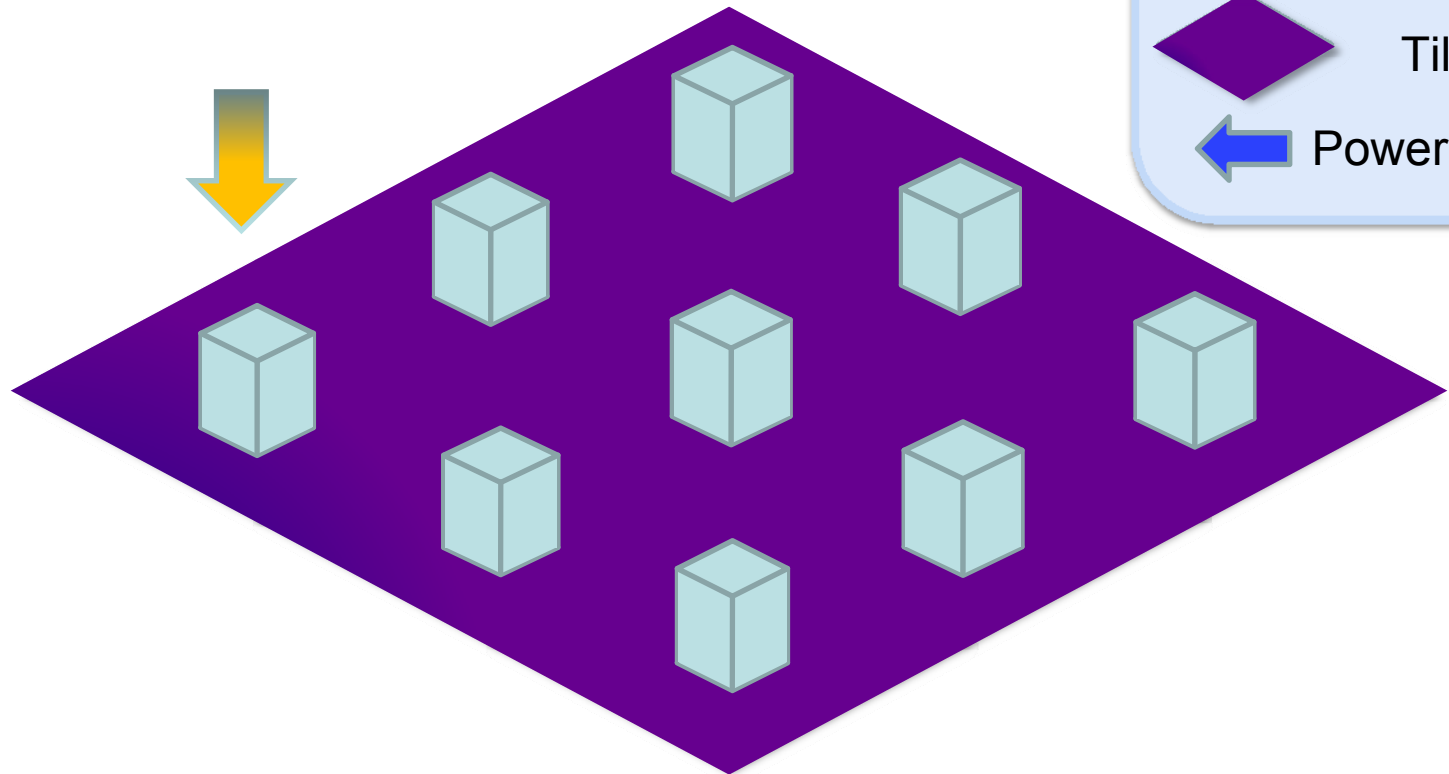
$$sell = sell - a_s \cdot temp$$

- Sell incentive must also consider running tasks:

$$sell = sell - \sum_{task_i} p_i$$

Agent-Based Power Trading

**Task is mapped to tile with
insufficient power units**



Agent-Based Power Trading

Agent trades power units with neighbors using cost function



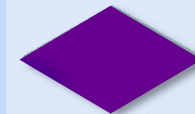
$$(sell_n - buy_n) - (sell_i - buy_i) > \varepsilon$$



Used units



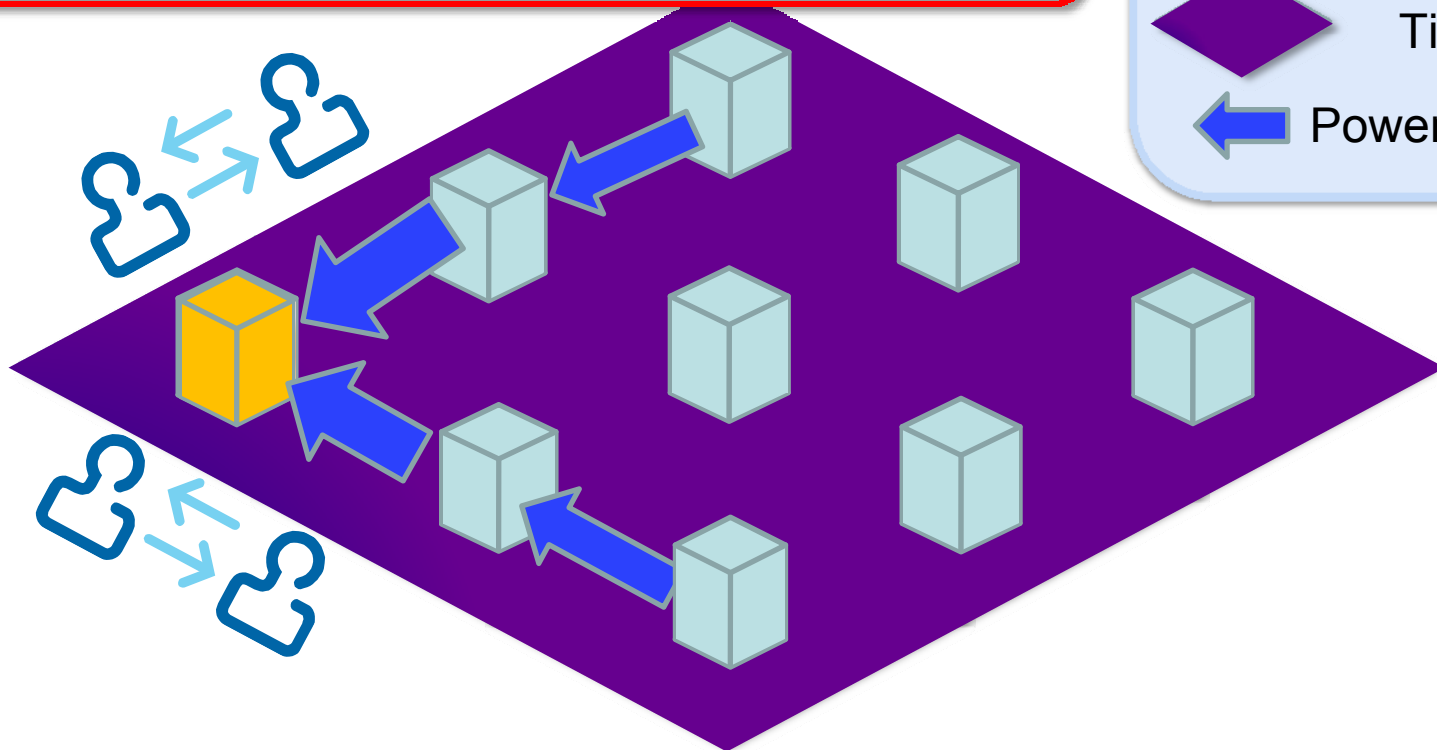
Free units



Tile

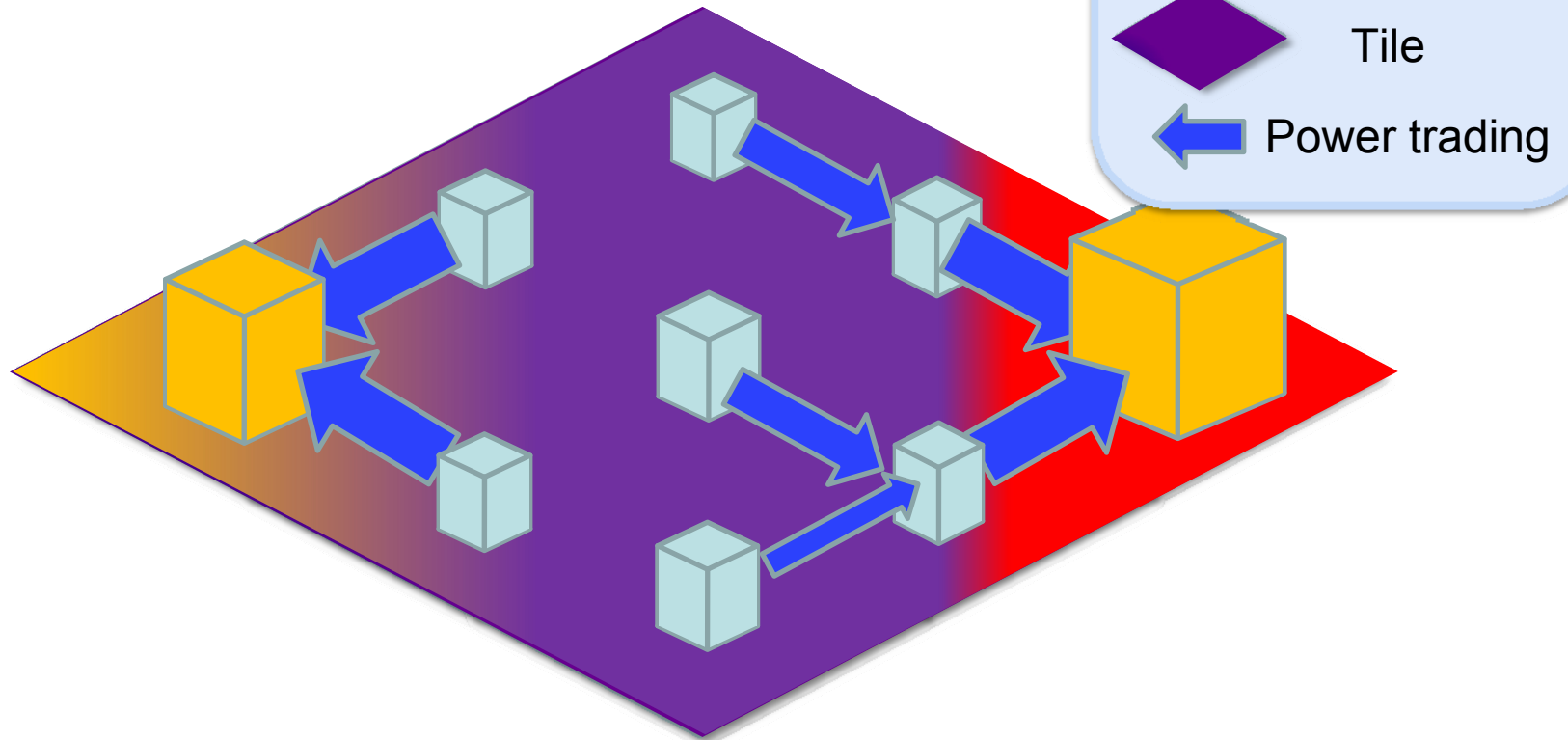


Power trading

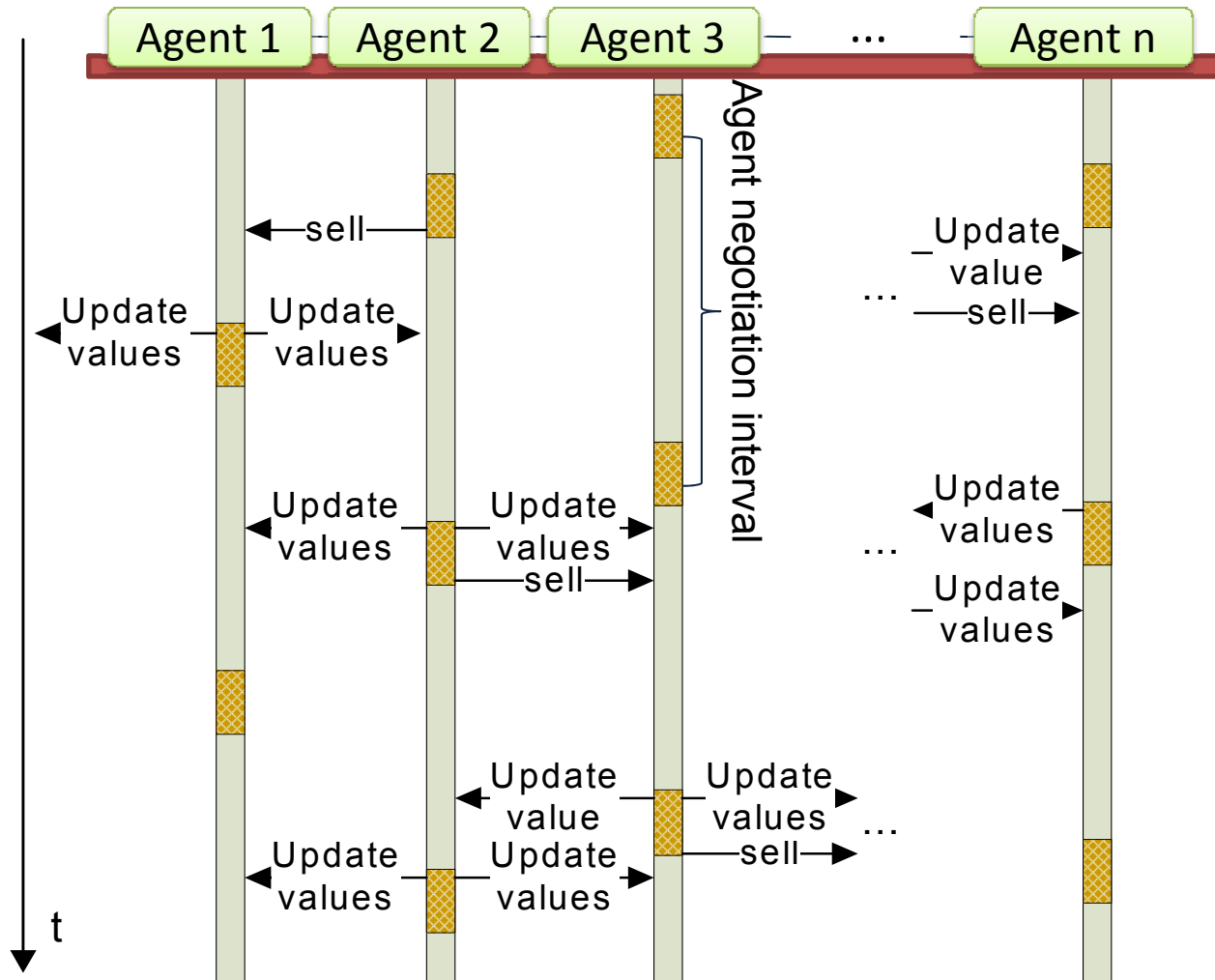


Agent-Based Power Trading

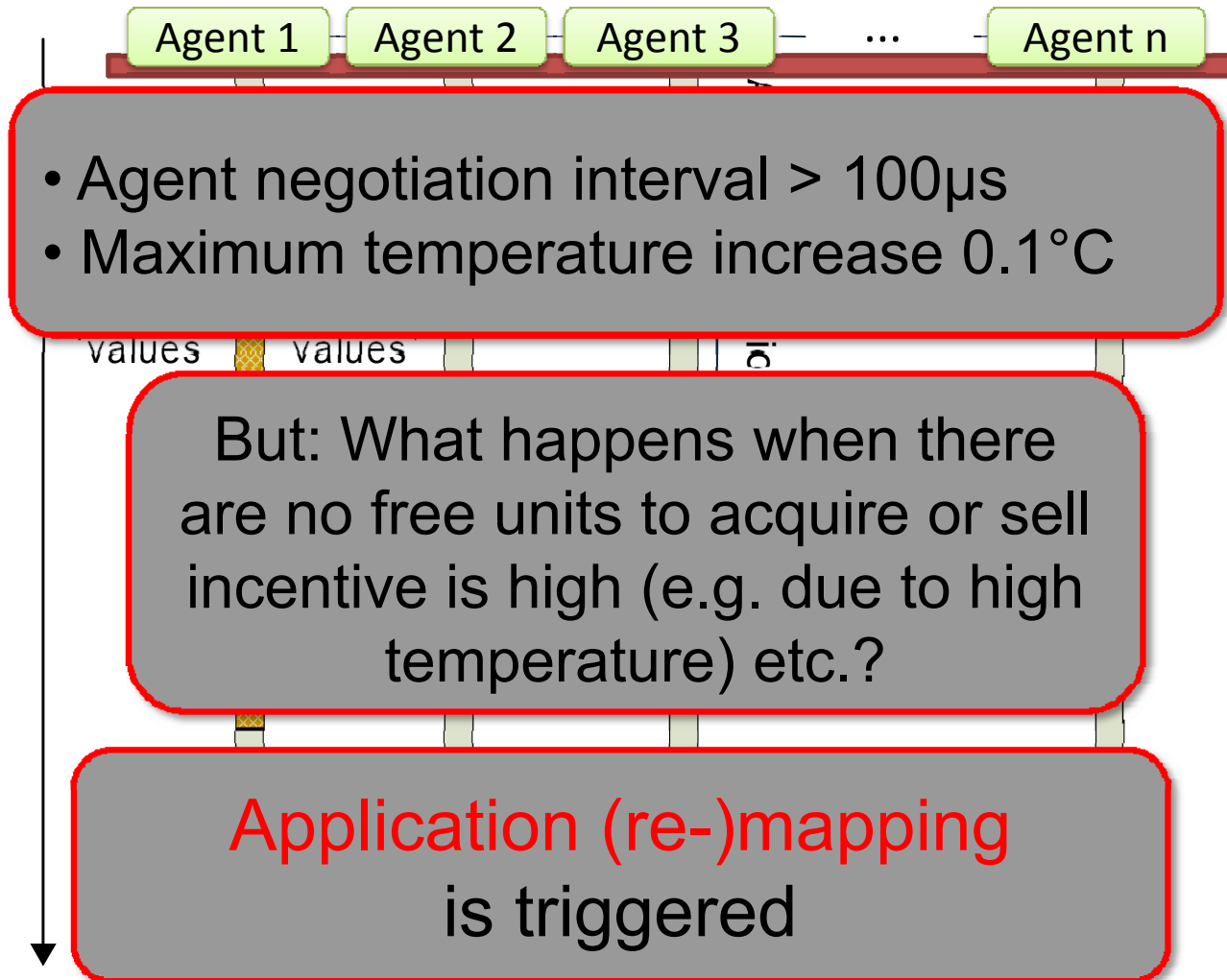
Temperature increase
may trigger new trading



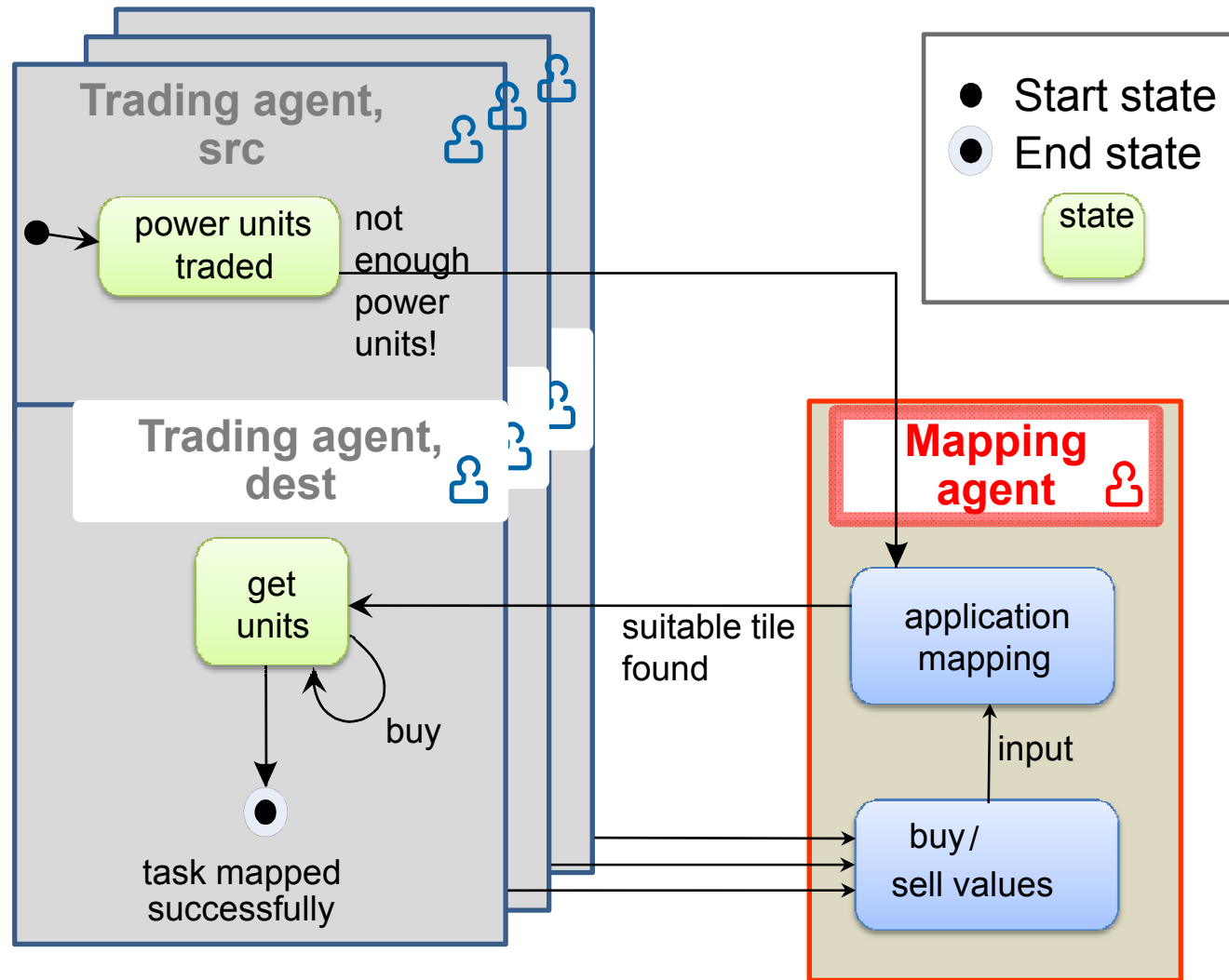
Power Unit Propagation



Power Unit Propagation



Task (Re-)Mapping



Task (Re-)Mapping

Triggered when tile does not have enough power units to run task

Mapping agents realized separately from trading agents

Buy and sell values input to mapping agents

Power trading agent of destination tile may require additional units

Task mapped to tile where $(sell_n - buy_n) - (sell_i - buy_i)$ is maximal as long as buy value is not negative.

Start state
End state

state

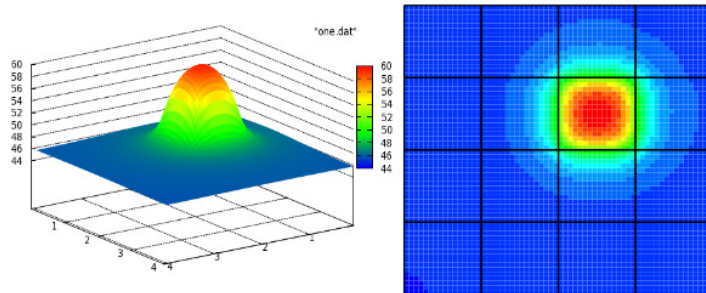
ing
nt



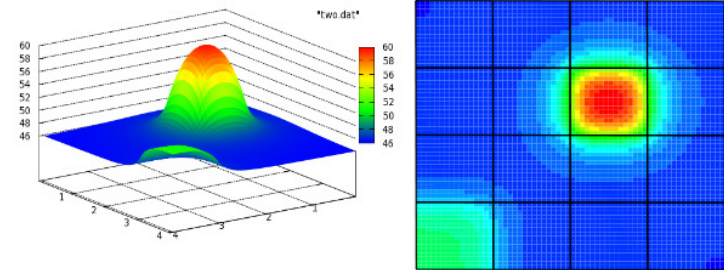
n

out

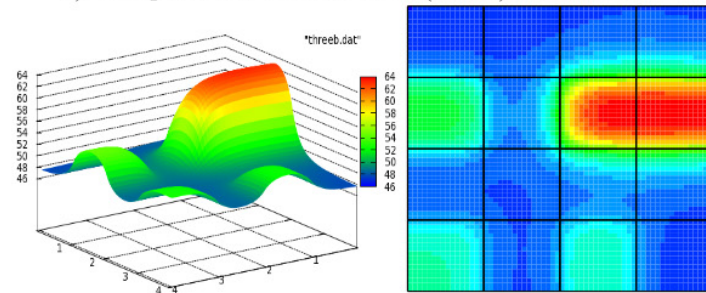
Agent-Based Power Trading Example



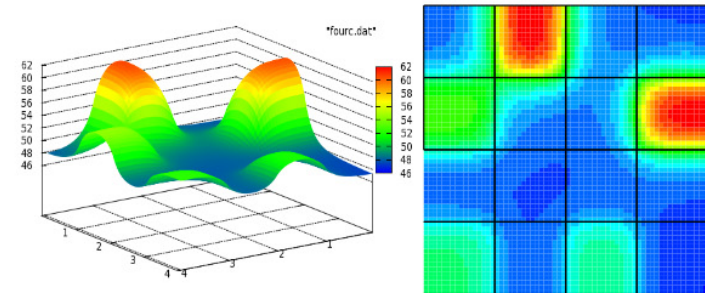
a) Temperature for one task (task₁) at time t_1



b) Temperature for two tasks at time t_2



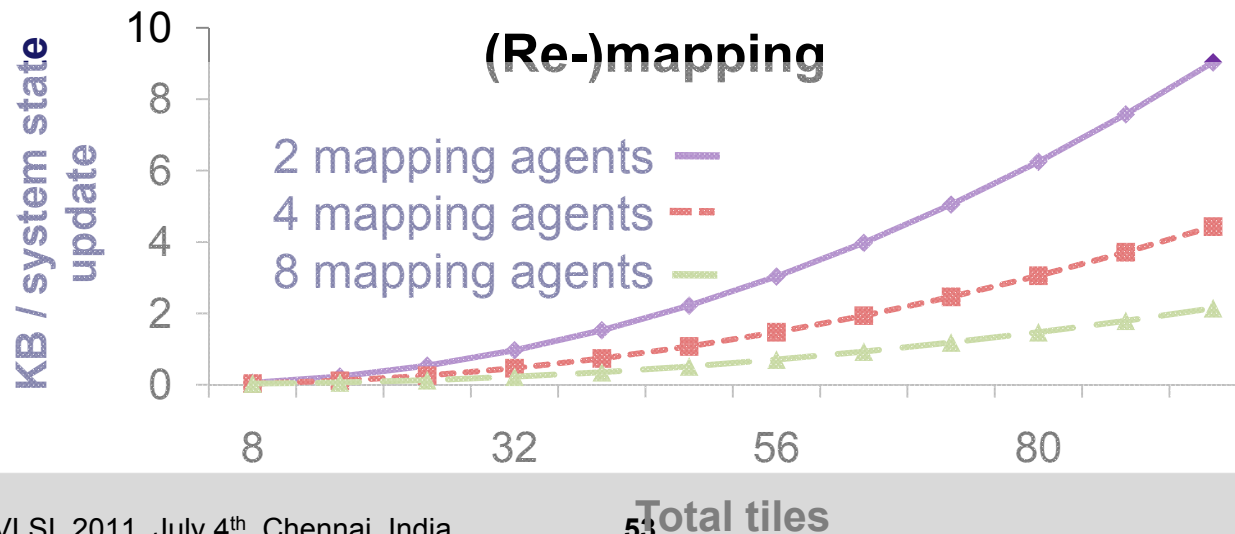
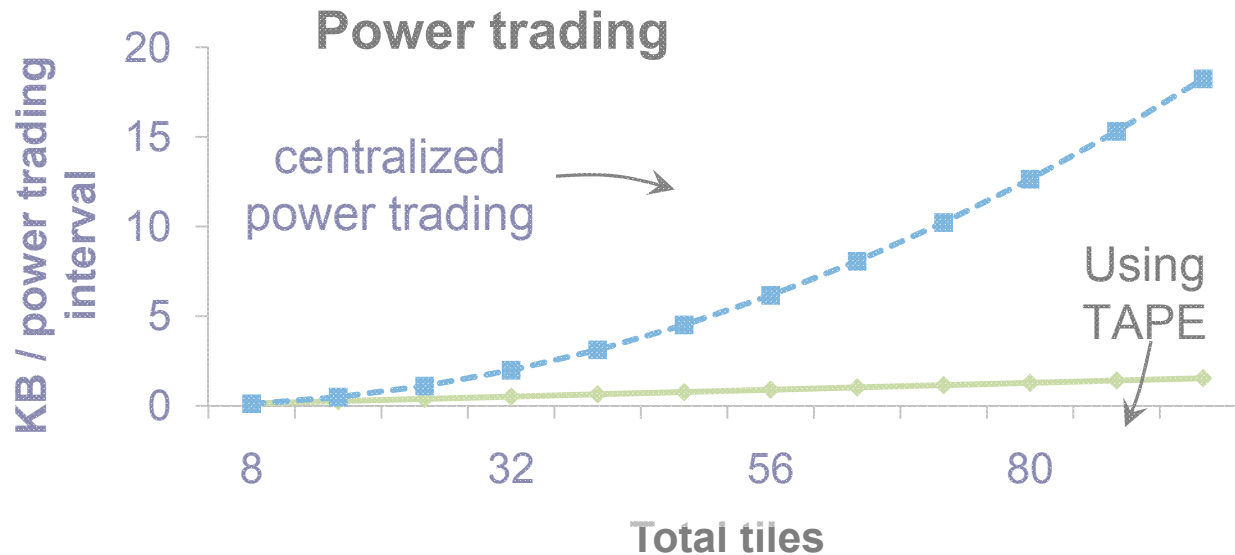
c) Five tasks at time t_i



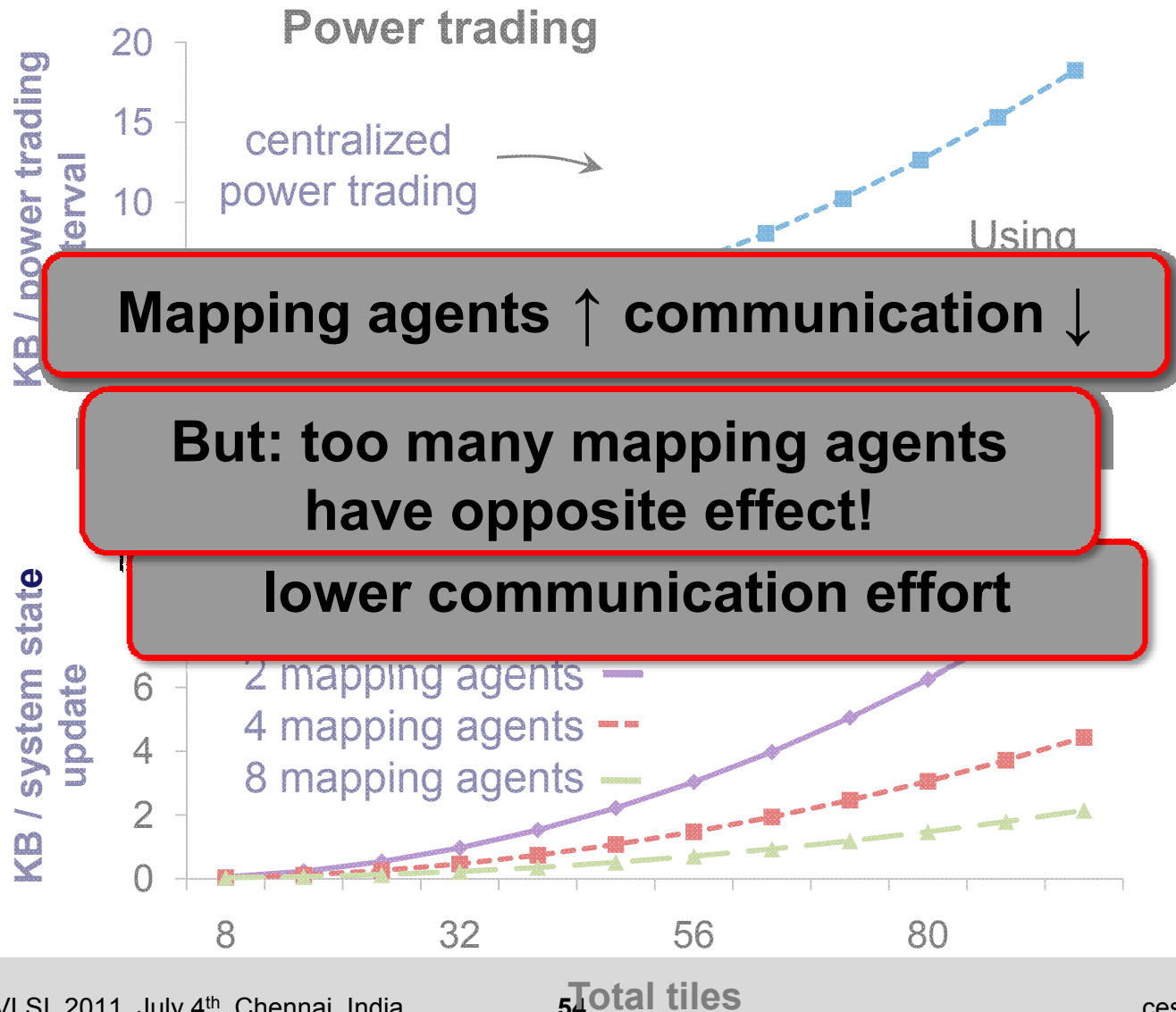
d) Task₁ (re-)mapped at time t_{i+1}

$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ <i>used</i>	$\begin{pmatrix} 0.5 & \dots & 0.5 \\ \vdots & \ddots & \vdots \\ 0.5 & \dots & 0.5 \end{pmatrix}$ <i>buy</i>	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ <i>used</i>	$\begin{pmatrix} 0.5 & 0.5 & 0.6 & 0.5 \\ 0.5 & 0.6 & 2.4 & 0.5 \\ 0.5 & 0.5 & 0.6 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 \end{pmatrix}$ <i>buy</i>	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ <i>used</i>	$\begin{pmatrix} 0.5 & 0.5 & 0.6 & 0.5 \\ 0.5 & 0.6 & 0.3 & 0.5 \\ 0.6 & 0.5 & 0.6 & 0.5 \\ 1.2 & 0.6 & 0.5 & 0.5 \end{pmatrix}$ <i>buy</i>	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ <i>used</i>	$\begin{pmatrix} 0.6 & 0.5 & 0.6 & 0.6 \\ 1.2 & 0.7 & -1.5 & 2.5 \\ 0.6 & 0.5 & 0.6 & 0.7 \\ 1.2 & 0.6 & 1.3 & 0.6 \end{pmatrix}$ <i>buy</i>	$\begin{pmatrix} 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ <i>used</i>	$\begin{pmatrix} 0.7 & 2.5 & 0.6 & 0.6 \\ 1.2 & 0.7 & 0.7 & 2.4 \\ 0.6 & 0.5 & 0.6 & 0.6 \\ 1.2 & 0.6 & 1.2 & 0.6 \end{pmatrix}$ <i>buy</i>		
$\begin{pmatrix} 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{pmatrix}$ <i>free</i>	$\begin{pmatrix} 0.6 & \dots & 0.6 \\ \vdots & \ddots & \vdots \\ 0.6 & \dots & 0.6 \end{pmatrix}$ <i>sell</i>	$\begin{pmatrix} 3 & 3 & 2 & 3 \\ 3 & 2 & 5 & 3 \\ 3 & 3 & 2 & 3 \\ 3 & 3 & 3 & 3 \end{pmatrix}$ <i>free</i>	$\begin{pmatrix} 0.6 & 0.6 & 0.4 & 0.6 \\ 0.6 & 0.4 & 2.2 & 0.6 \\ 0.6 & 0.6 & 0.4 & 0.6 \\ 0.6 & 0.6 & 0.6 & 0.6 \end{pmatrix}$ <i>sell</i>	$\begin{pmatrix} 3 & 3 & 2 & 3 \\ 3 & 2 & 5 & 3 \\ 2 & 3 & 2 & 3 \\ 4 & 2 & 3 & 3 \end{pmatrix}$ <i>free</i>	$\begin{pmatrix} 0.6 & 0.6 & 0.4 & 0.6 \\ 0.6 & 0.4 & 2.2 & 0.6 \\ 0.4 & 0.6 & 0.4 & 0.6 \\ 1.4 & 0.4 & 0.6 & 0.6 \end{pmatrix}$ <i>sell</i>	$\begin{pmatrix} 2 & 3 & 2 & 2 \\ 4 & 1 & 5 & 4 \\ 2 & 3 & 2 & 1 \\ 4 & 2 & 3 & 2 \end{pmatrix}$ <i>free</i>	$\begin{pmatrix} 0.4 & 0.6 & 0.4 & 0.4 \\ 1.4 & 0.3 & -1.1 & 1.9 \\ 0.4 & 0.6 & 0.4 & 0.3 \\ 1.4 & 0.4 & 1.3 & 0.4 \end{pmatrix}$ <i>sell</i>	$\begin{pmatrix} 1 & 4 & 2 & 2 \\ 4 & 2 & 2 & 5 \\ 2 & 3 & 2 & 2 \\ 4 & 2 & 4 & 2 \end{pmatrix}$ <i>free</i>	$\begin{pmatrix} 0.3 & 1.9 & 0.4 & 0.4 \\ 1.2 & 0.4 & 0.4 & 2.0 \\ 0.4 & 0.6 & 0.4 & 0.4 \\ 1.2 & 0.4 & 1.2 & 0.4 \end{pmatrix}$ <i>sell</i>		
$\underbrace{\hspace{10em}}$ <i>time</i> t_0		$\underbrace{\hspace{10em}}$ t_1		$\underbrace{\hspace{10em}}$ t_2		\dots		$\underbrace{\hspace{10em}}$ t_i		$\underbrace{\hspace{10em}}$ t_{i+1}	

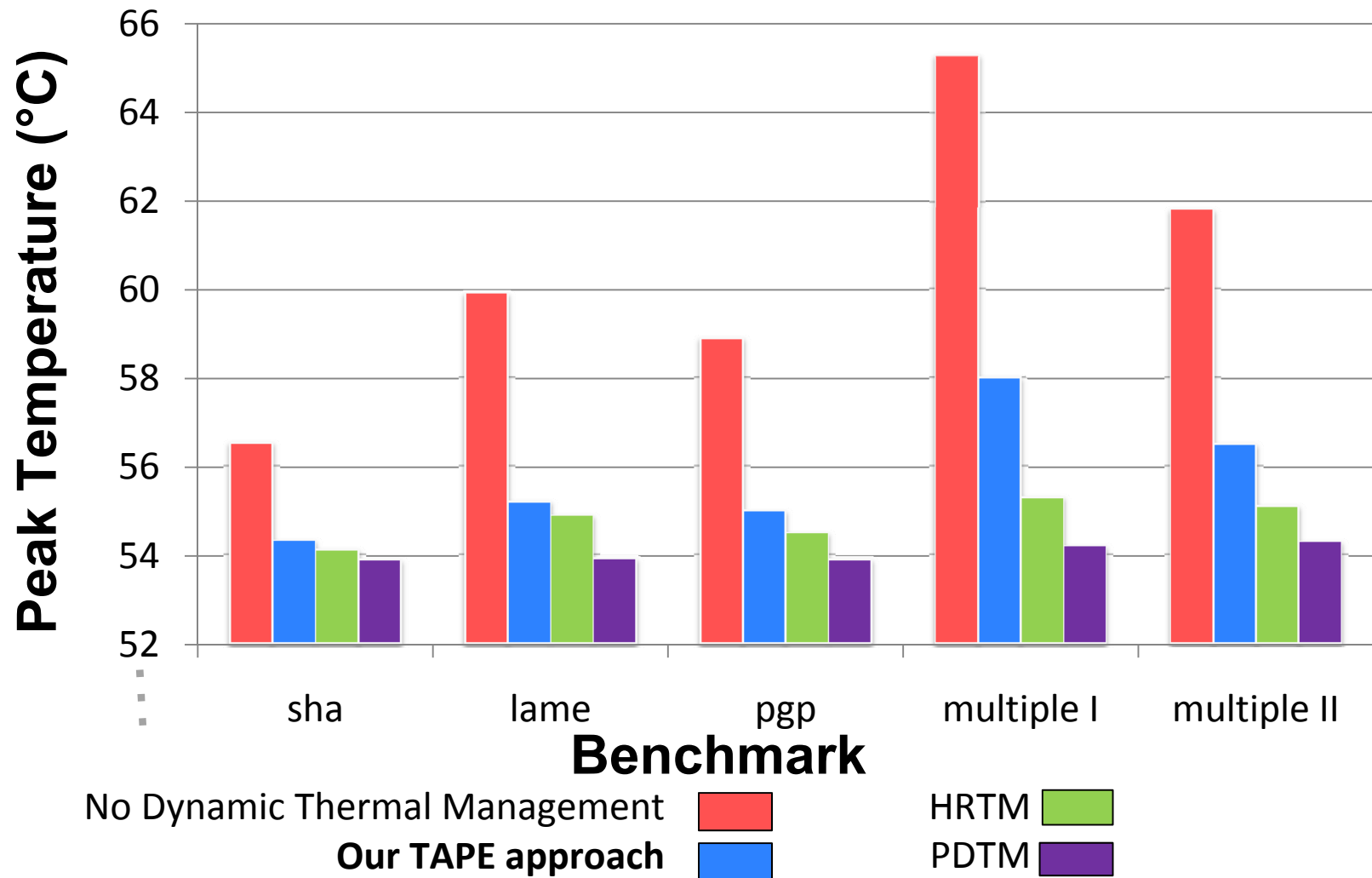
Scalability of Agent Communication



Scalability of Agent Communication



Results: Peak Temperature



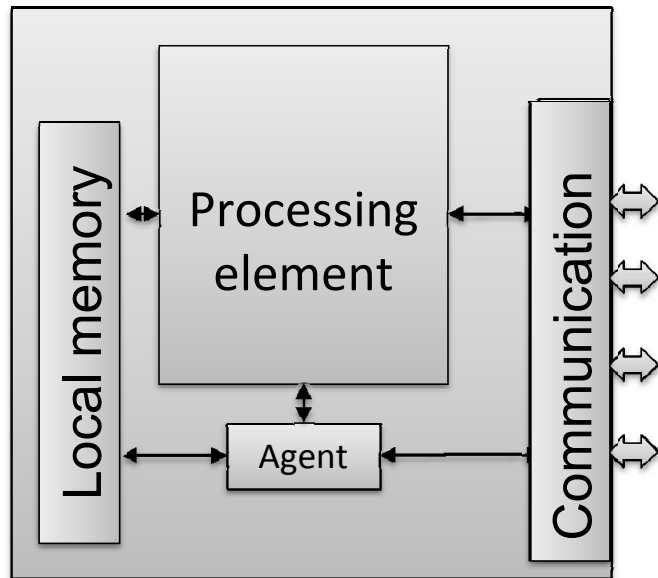
No Dynamic Thermal Management

Our TAPE approach

HRTM

PDTM

Agent Implementations

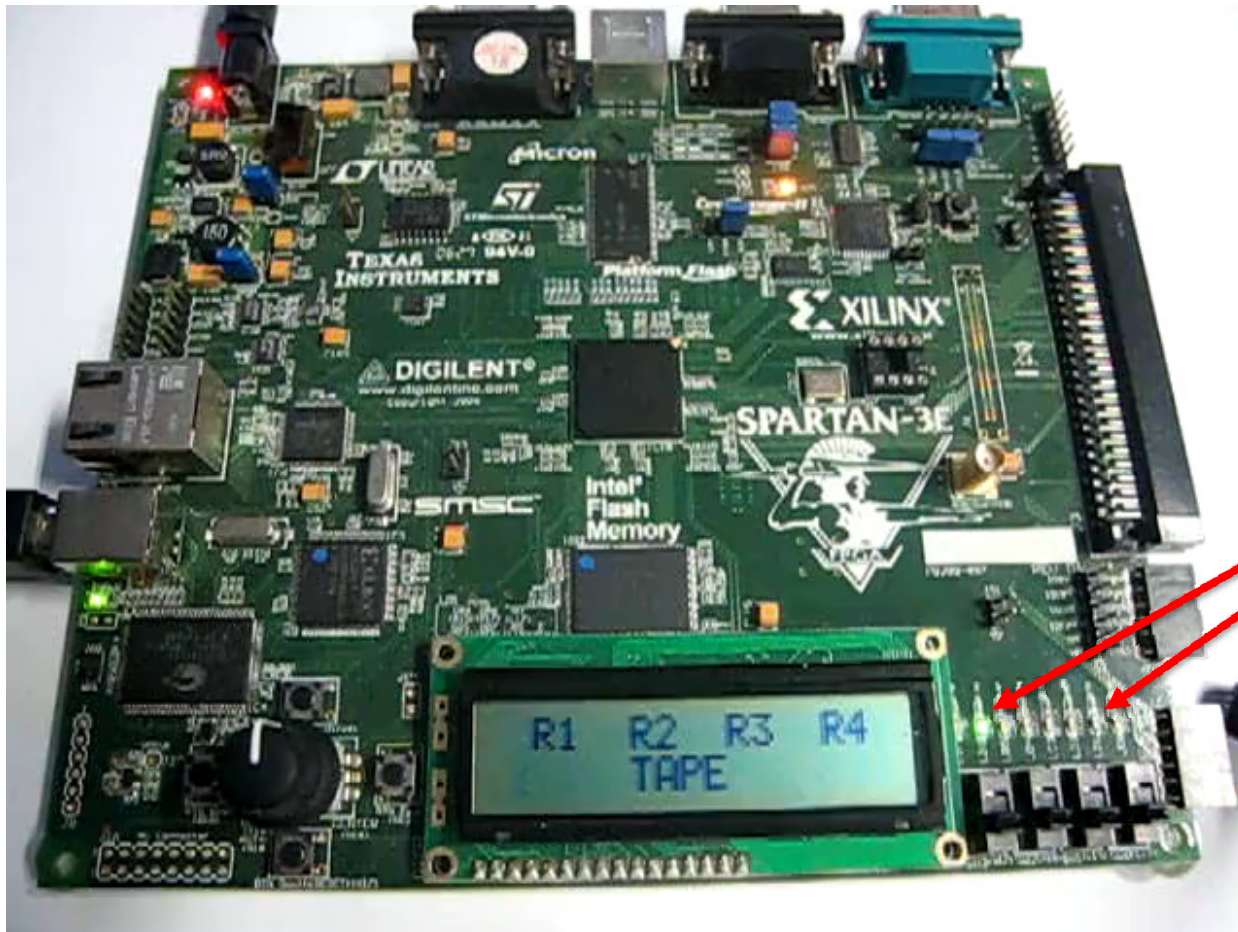


- Implemented in Software
 - Compete with tasks for computation
 - Are not always possible (e.g. in dedicated hardware)
- Implemented in Hardware
 - Can be realized on any tile
 - Does not take processing time away from tasks
 - Require additional area (143 slices in Xilinx Virtex-4 FPGA)

slices	LUTs	Flip-flops	Mult	Max Freq
143	276	84	2	148.9MHz

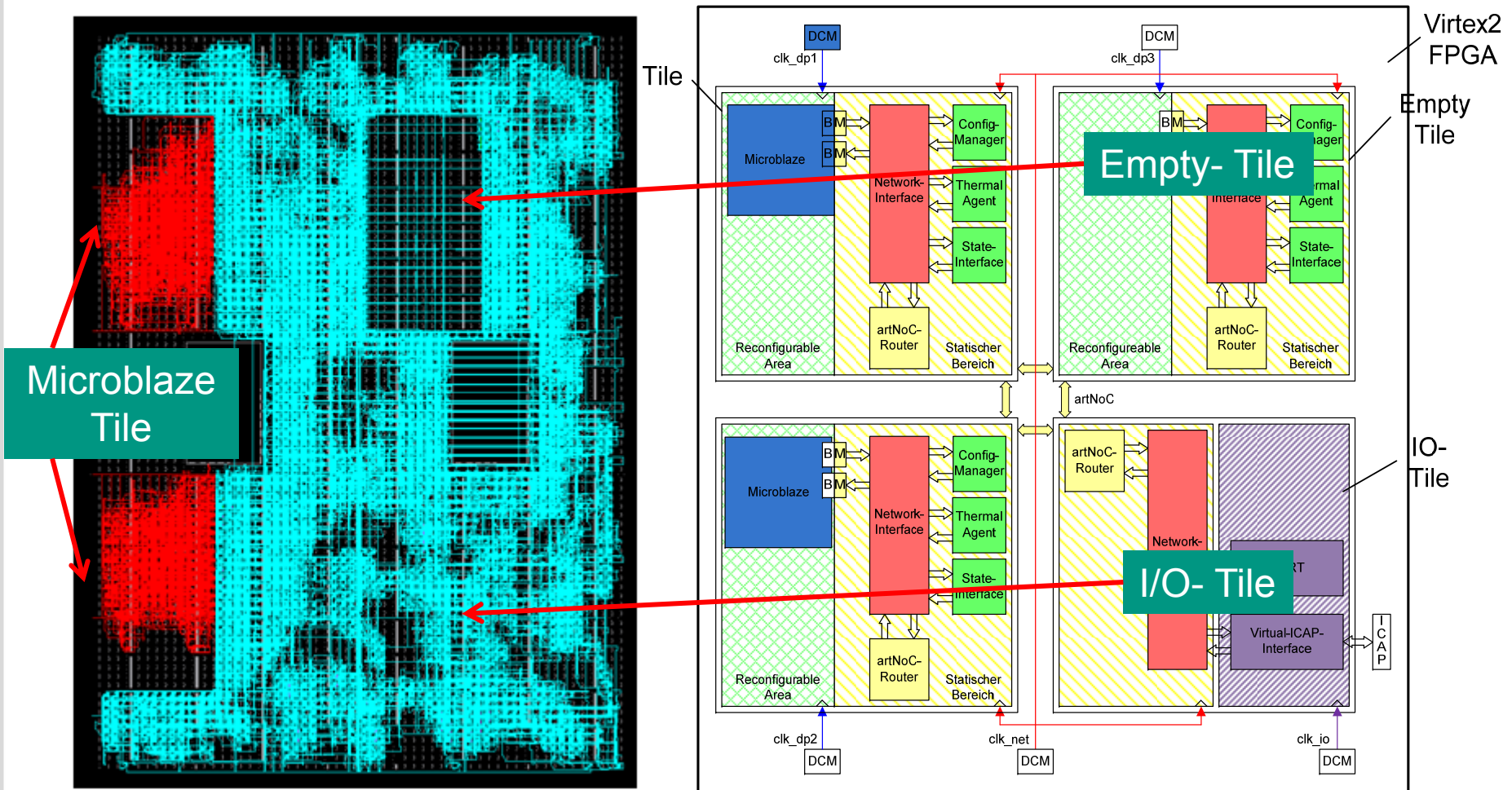
Hardware Demonstrator

- Hardware prototype in running on a Xilinx Spartan3e FPGA with 4 Picoblaze tiles
- Thermal sensors realized through ring oscillators

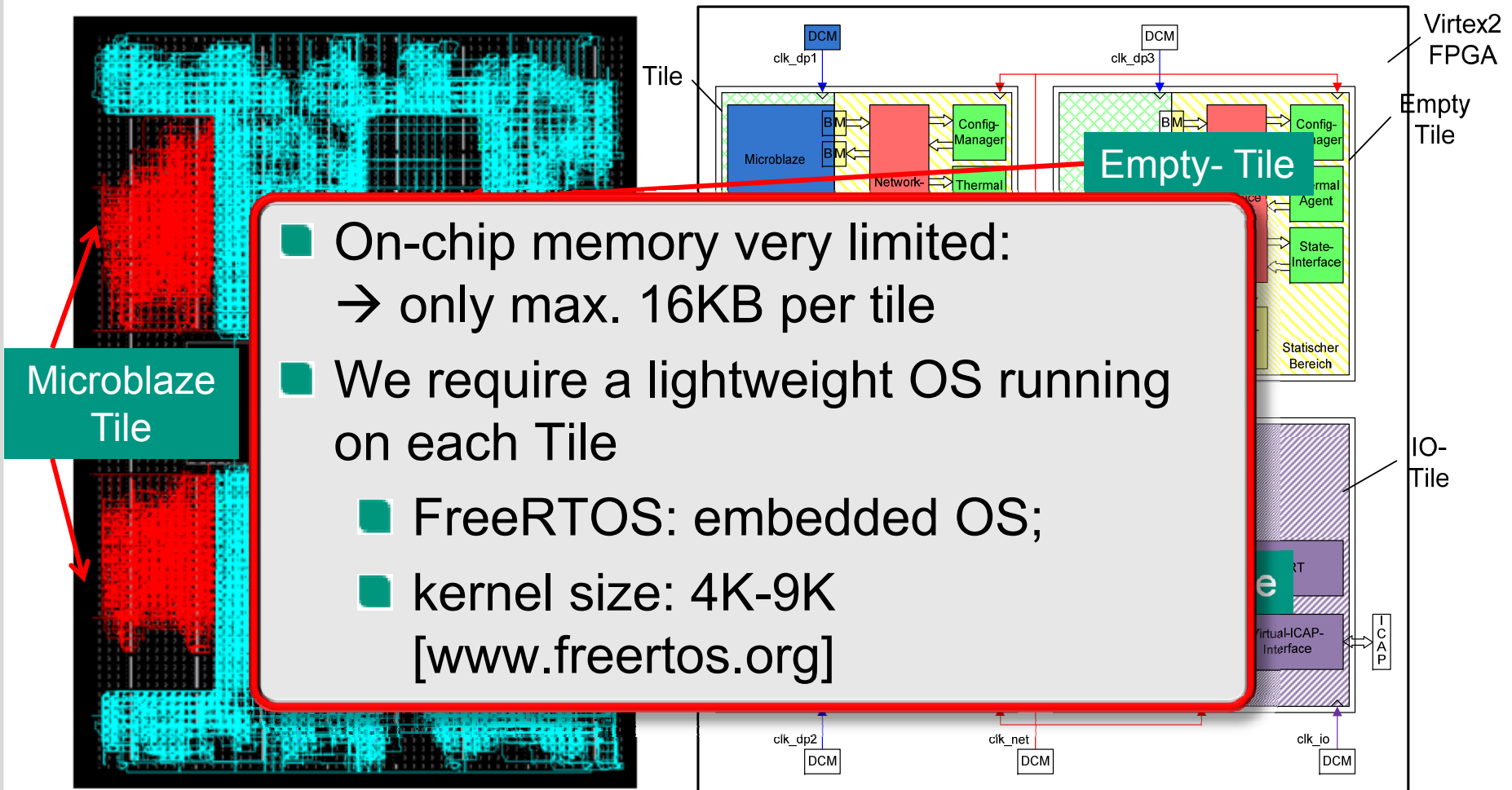


Watch LEDs

System Setup

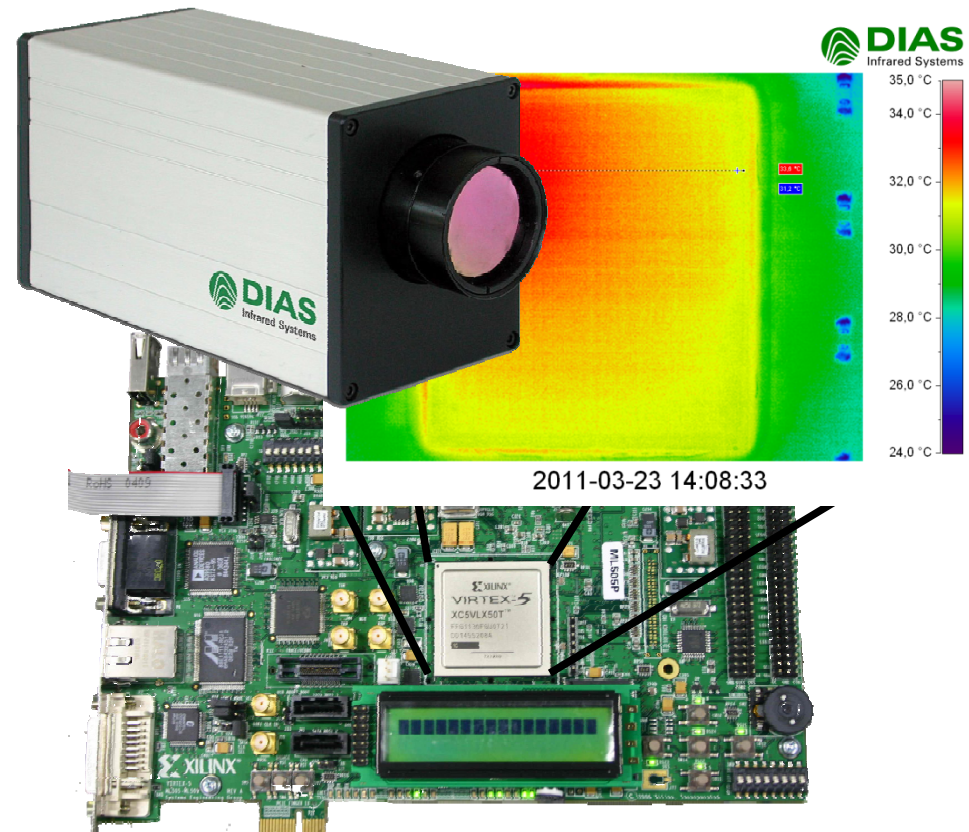


System Setup

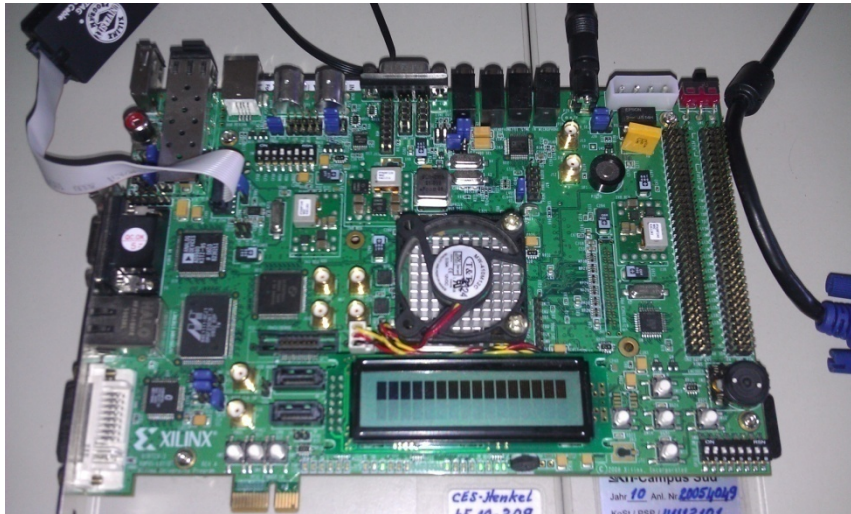


Thermal Camera for accurate thermal Evaluation

- DIAS Pyroview IR Camera
 - Spatial resolution macro lens: around 50 μ m
 - Limited by camera IR spectral range of 8 μ m- 14 μ m
 - Temperature range configurable -20 °C to 120 °C or 0°C to 500°C
 - Sampling rate of 50Hz
 - Camera transmits 50 frames per second over ethernet in real time
 - 384x288 pixels
 - Comprehensive SDK for accessing camera functionality



System Setup



Summary

- Reliability is a problem when migrating to upcoming technology nodes
- MTTF of certain effects are related to temperature
- Dynamic Thermal Management techniques are necessary
- Important features:
 - Scalability -> for many core systems with 100s of cores
 - Single point of failure for DTM should be avoided
- Principles of self-organization may be a solution



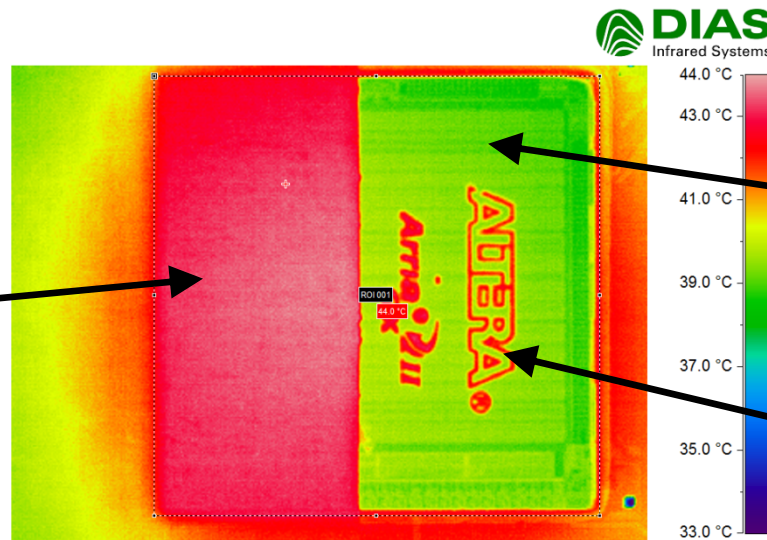
The background of the slide is a photograph of the KIT main building, a large, historic, yellow-colored structure with multiple wings and a central tower. The building is situated behind a large, green lawn. In the foreground, there is a gravel path and a low, dark stone wall. The sky is blue with scattered white clouds. The text "Thank you for Attention!" is overlaid in the center of the image in a large, white, sans-serif font.

Thank you for Attention!

Infrared Measurements and Emissivity

- Emissivity can be a problem for infrared measurements
 - Ideal „black body“ has emissivity of 1
 - Polished metal can be as low as 0.01
 - Emissivity of Silica: 0.9 - relatively high, but not optimal
- Low emissivity results in high reflection of surrounding temperatures

Masking tape
(emissivity 0.95)
covering half of
chip shows actual
temperature

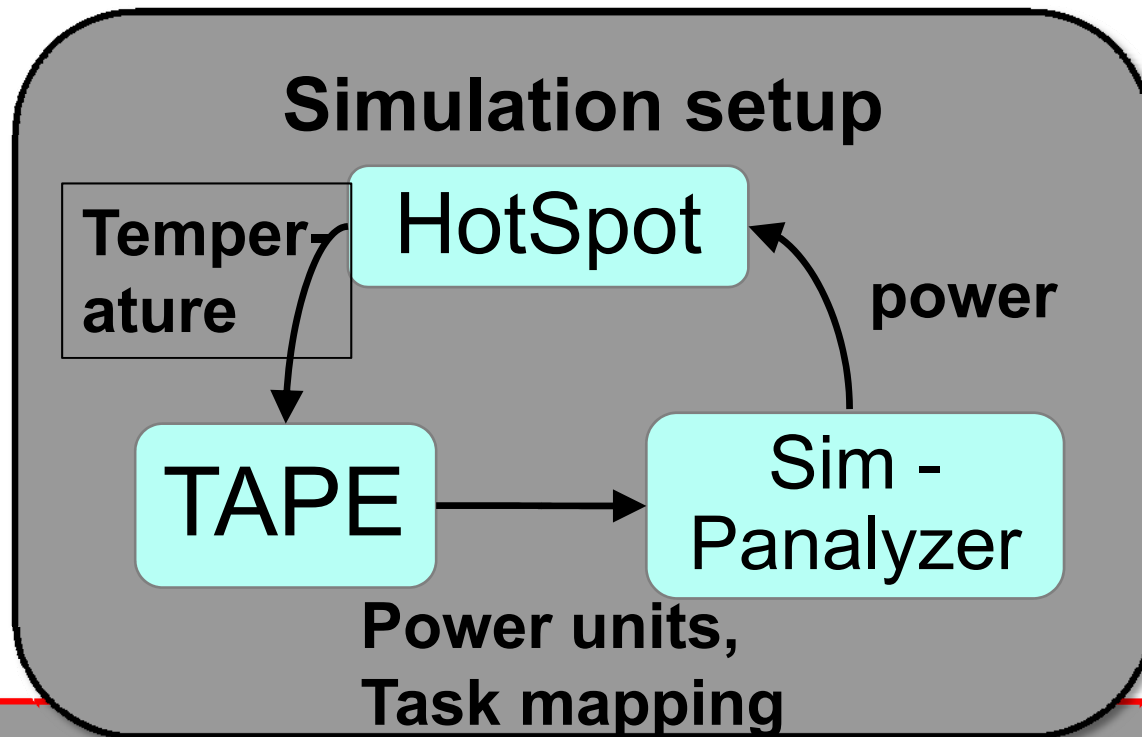


Metal packaging
measurements very
inaccurate

Paint of logo has
high emissivity
(around 0.92)

[Emissivity table of various materials: www.omega.com/temperature/z/pdf/z088-089.pdf]

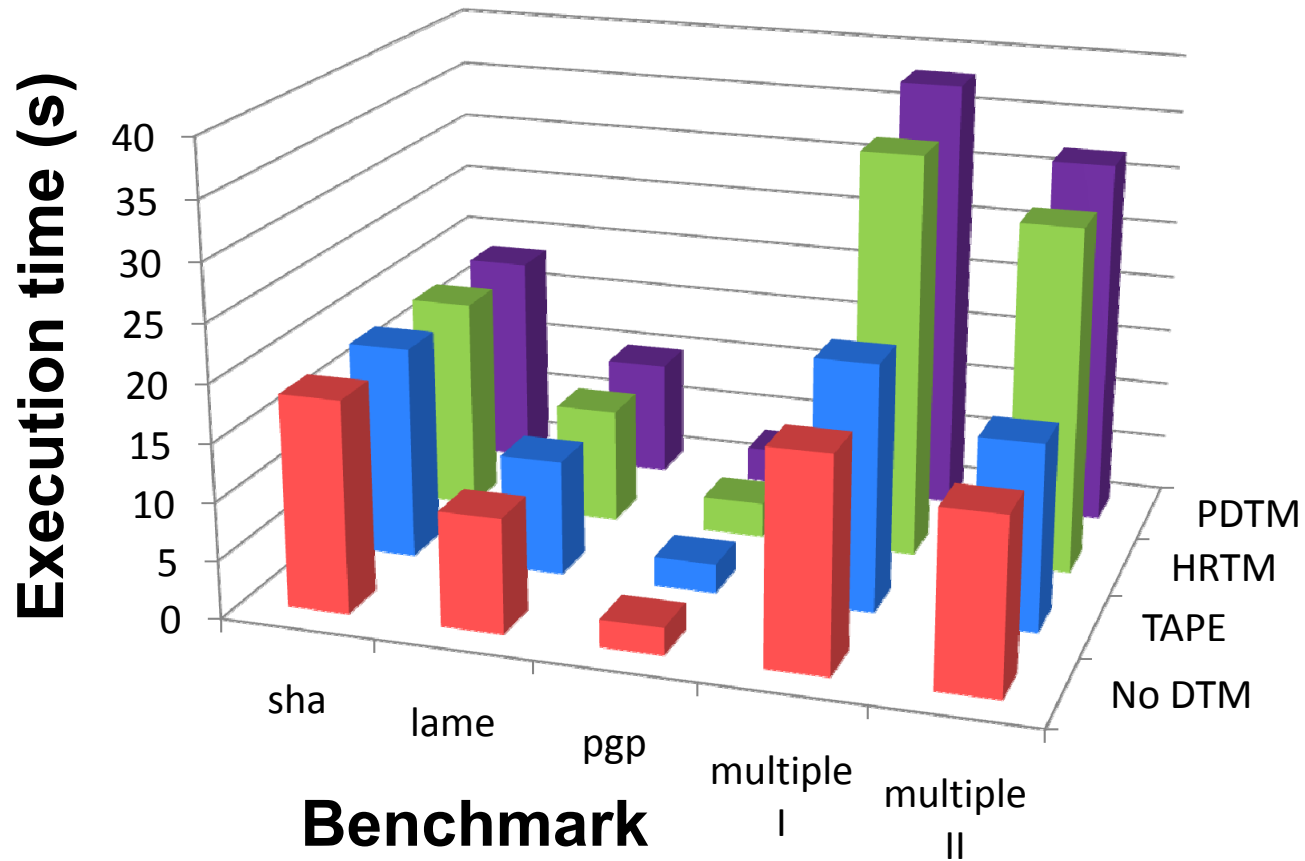
Results: Execution Time



Dynamic temperature threshold greatly reduces execution time (44%) due to less frequent task migrations

Task migration penalty is 100K cycles (saving, transferring, loading task context)

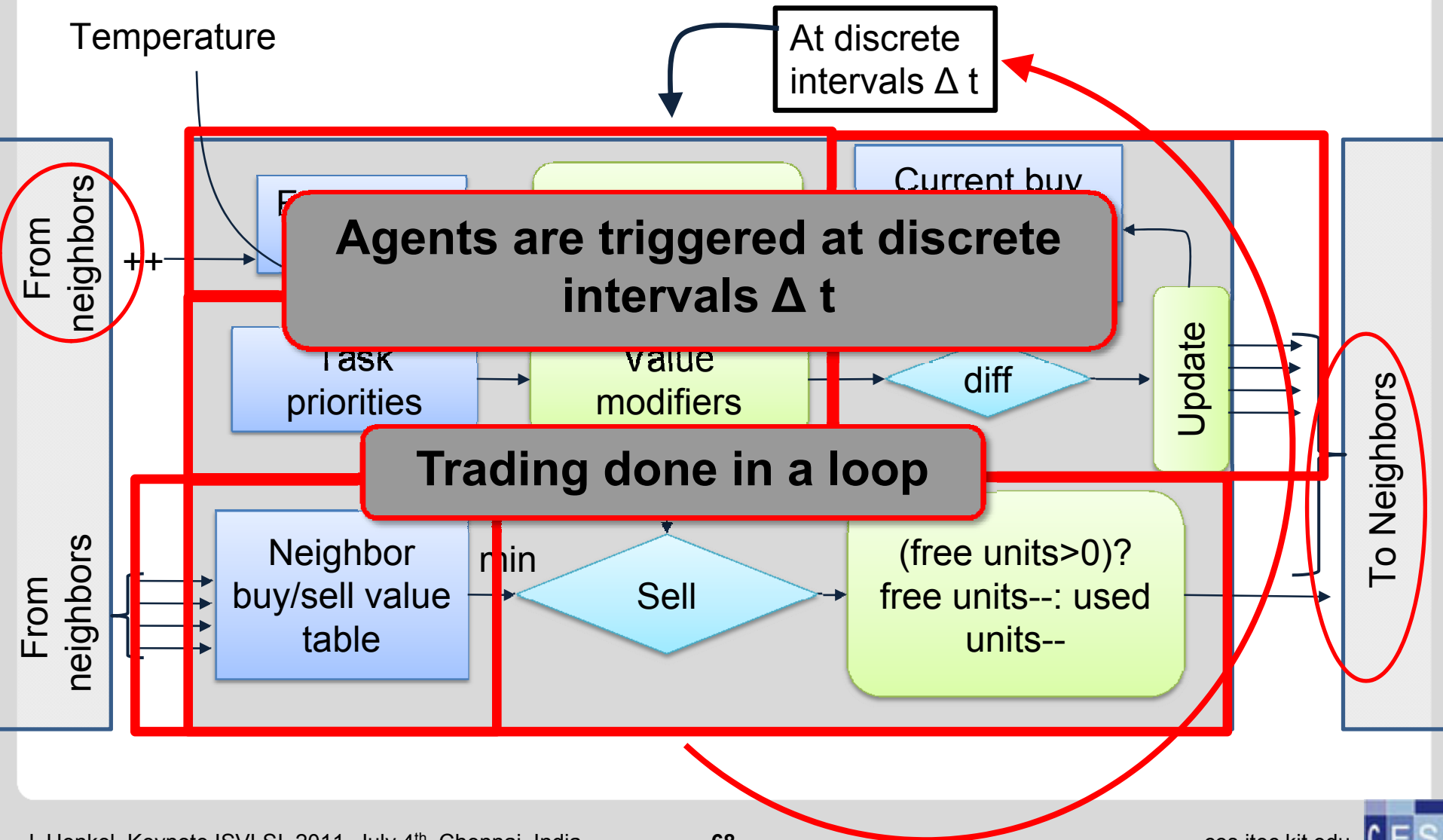
Results: Execution Time



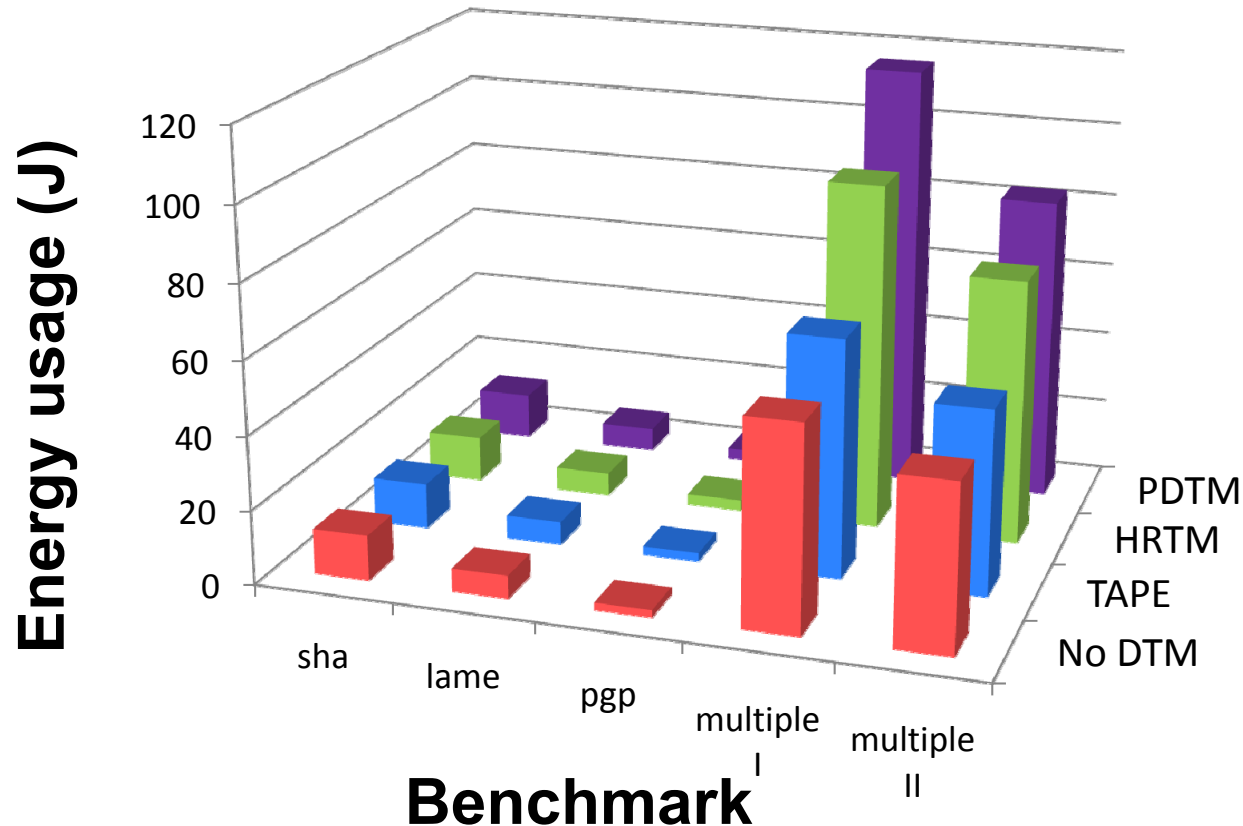
No Dynamic Thermal Management
Our TAPE approach

HRTM
PDTM

Agent-Based Power Trading



Results: Total Energy Consumption



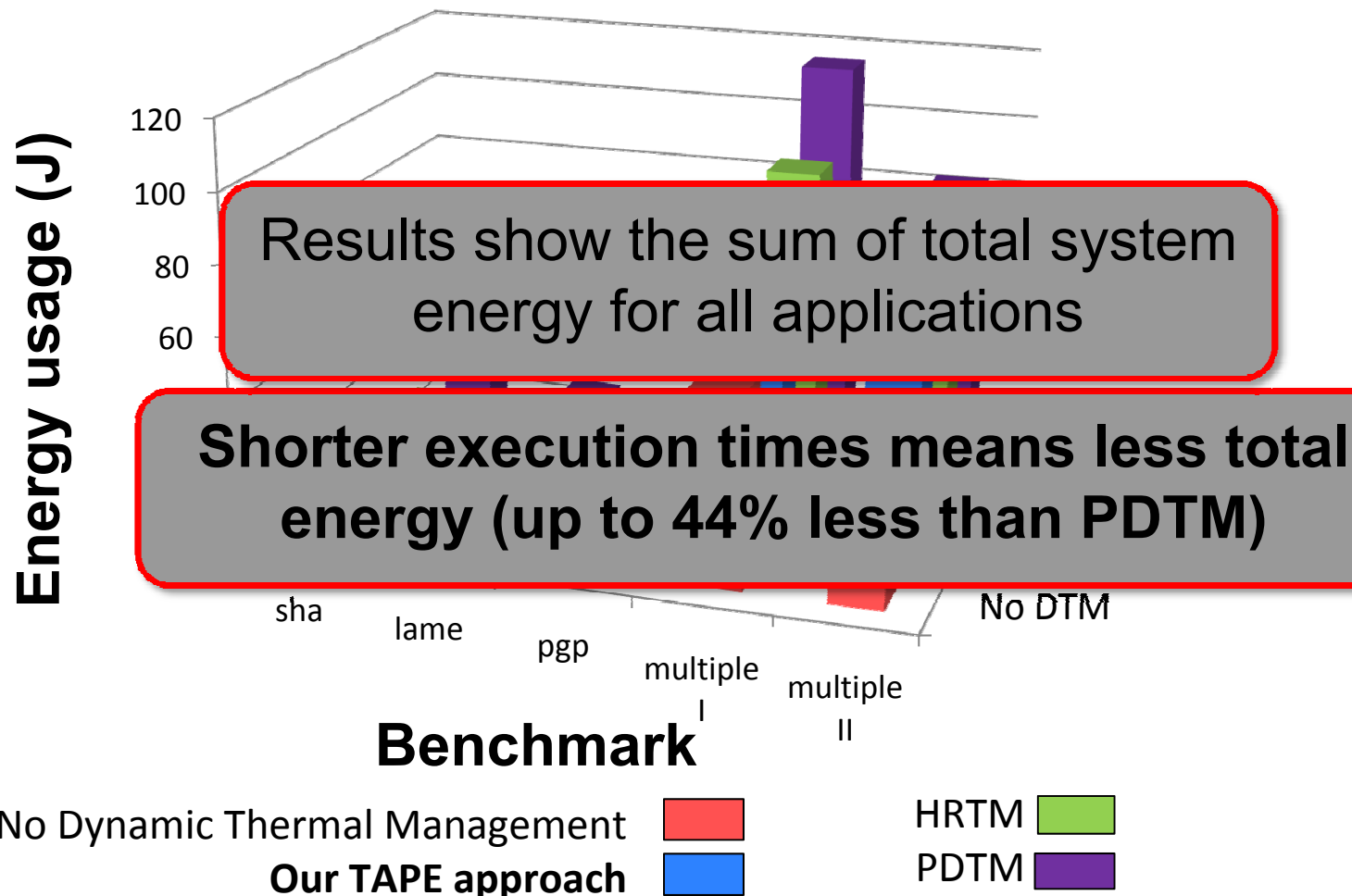
No Dynamic Thermal Management

Our TAPE approach

HRTM

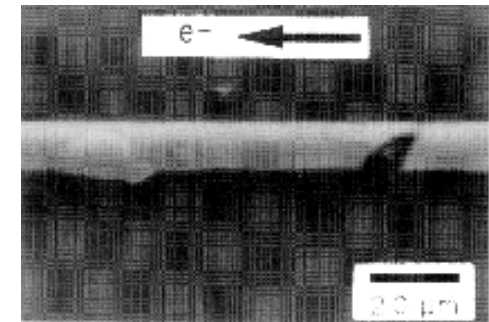
PDTM

Results: Total Energy Consumption

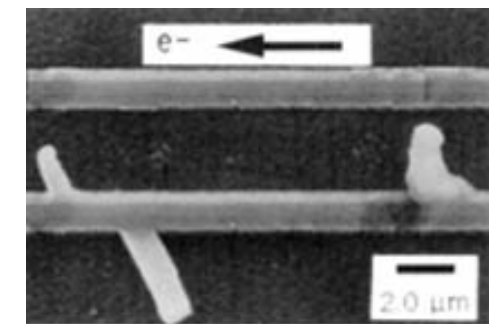


Temperature and Reliability

- Process variations and electromigration can result in **hillocks** and **holes**
 - Lead to open failures or short circuit failures respectively
 - Failures may be temperature dependent due to material expansion
 - Holes may function normally at high temperatures but fail at low temperatures
 - Hillocks may function normally at low temperatures but short circuit at high temperatures



Hole/crack



Hillock

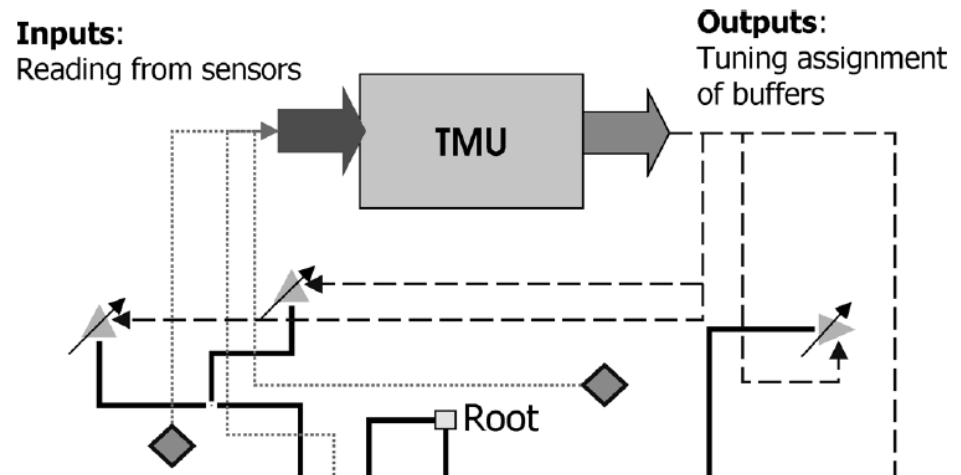
[W.D. Nix, 1992]

ces.itec.kit.edu

Temperature and Reliability

- Transient errors may result due to timing errors
 - Approx. 5% decrease in delay every 10°C temperature increase [Xie 2006]
 - Timing errors result from spatial temperature variations
 - localized **hotspots** need to be avoided
 - Clock trees are particularly vulnerable
 - Span across multiple thermal areas
 - Additional buffers can be inserted to cope with thermal clock skew

Clock skew compensation using a **thermal management unit** to control **tunable delay buffers** inserted into clock tree



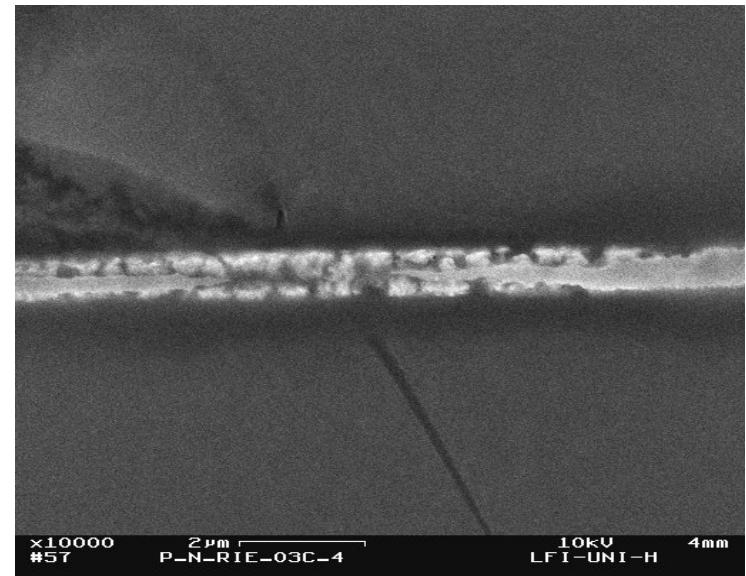
Temperature and Reliability

- **Electromigration**: aging effect due to transport of mass in metal interconnects
- directly linked to temperature
 - Basic *Mean time to failure* modeled by Black's Equation:

$$MTTF = Aj^{-n} e^{\left(\frac{Q}{kT}\right)}$$

- *MTTF* decreases exponentially with temperature

→ **Goal**: reduce peak temperatures



[wikipedia]

Thermal issues in 3D

- Power density increases with technology scaling
 - On an average from 2W/mm^2 in 65nm technology to 7.2 W/mm^2 for 45nm [Vijaykrishnan et al. ISQED'06]
- Higher power density and temperature variation cause transient and permanent failures
 - Due to technology scaling, a drop of 66% in feature size increases the temperature from 342°K to 356°K and reduces the MTTF by 76% [Srinivasan et al. Micro'05]
- Leakage power increases with temperature
 - A change in temperature from 40°C to 120°C increases the leakage power 4 times [Li et al. NOCS'08]

Categorization of technology induced dependability effects

■ I. Process and design time effects

- Yield and process variability
- Complexity: $> 10^{11}$ (100 billion) within a decade

■ II. **Operation and run-time effects**

- Aging effects (irreversible)
- **Thermal effects** (may speed up some ageing effects)
 - Aggressive power management may be counter-productive since thermal cycling is increased -> tradeoff
- Soft errors
 - $> 8\%$ increase per technology node
 - Errors are random and transient and limit exploitation of techniques like voltage scaling

Counter Measures (cont'd)

*“Emerging devices are expected to be **more defective**, **less reliable** and **less controlled** in both their position and physical properties.”*

“It is therefore important to go beyond simply developing fault-tolerant systems that monitor the device at run-time and react to error detection.”

*“It will be necessary to **consider error as a specific design constraint** and to develop methodologies for error resiliency, accepting that **error is inevitable** and trading off error rate against performance (e.g. speed, power consumption) in an application-dependent manner.”*

Source ENIAC Strategic Research Agenda, European Technology Platform Nanoelectronics ,2nd Ed., Nov. 2007.

Counter Measures (cont'd)

Some further citations:

- „ ... *build dependable systems with non-dependable components*“ (Shekhar Borkar, Intel, at IEEE/ACM DATE'07 Conference, Nice, 2007).
- Leon Stok, IBM: „... *most variability had been hidden from the designers ...This practice no longer holds for current [and future (Anm.)] technology nodes*“ (see p. 344 [D&T-JA08]).
- Jan Rabaey, UC Berkeley and Sharad Malik, Princeton: „*Existing solutions are unlikely to scale, and we will need radically new solutions ...*“ (see p. 299 [D&T-JA08]).
- [ITRS] on overall design technology challenges: “*Design Productivity* ” , “ *Power Consumption* ” , “ *Reliability* ” , “*Interference and Manufacturability*”

Temperature and Reliability

Black's equation

$$MTTF = A\omega j^{-n} e^{\left(\frac{Q}{kT}\right)}$$

Transient errors

Narayanan, V. and Xie, Y. *Reliability Concerns in Embedded System Designs. Computer* 39, 1 (Jan. 2006), 118-120.